Author:
Stevens, Sophie C C

Title:
Incidence Geometry in the Plane and Applications to Arithmetic Combinatorics

Incidence Geometry in the Plane and Applications to Arithmetic Combinatorics

Sophie Clara Charlotte Stevens

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Science.

School of Mathematics, 12th December 2019

Word count: 34,000

Abstract

In this thesis we study the applications of incidence geometry to additive combinatorics, with a particular focus of the setting of the prime residue field.

We begin by introducing some context to incidence geometry and its role in combinatorics. In Chapters 4 and 5, we prove a new incidence bound between points and lines in the plane and discuss its applications. The sum-product phenomenon is a motivating problem of this thesis and we discuss this in both the case of an arbitrary field and the specific case of $\mathbb{F} = \mathbb{R}$. In this latter setting, in Chapter 6 we prove a bound on the energy formulation of the sum-product problem. A consequence of this is to the sum-product problem itself, as well as to that of expander functions. Finally, in Chapter 8, we prove a new bound on the number of pinned distances in the plane.

Acknowledgements

This document could not and did not happen in isolation and this page presents a chance to begin to express my thanks towards those who have helped me produce it.

First and foremost I am indebted to my supervisor Misha Rudnev, for his wisdom, his patience and his kindness. Thank you to Lynne Walling for her accuracy in recommending him as a Good Person, and also for encouraging me to do a PhD.

Thanks to Julia Wolf, who provided me with countless opportunities, and to Olly Roche-Newton and Sean Prendiville who are doubtless unaware of the encouragement they gave me. I am grateful of the opportunities that the HARICOT and SPACE seminars provided, and to the combinatorics department at Bristol who provided the material and companionship.

I have learnt lots from working with Brendan Murphy, and also with Frank de Zeeuw and Ilya Shkredov. I am grateful to Jozsef Solymosi, especially for the hospitality he provided at the University of British Columbia.

My time at Bristol has been greatly enhanced by the people I have met who have kept me sane and happy. Ana Costache, Nick Jones and Luka Rimanić are the three who have played the greatest role in this task, but Justin D., Charley, Emma, Jenny, Simon, Lydia and Olly K. have all played their parts in my contentment.

For those I have have known the longest and who have given me support in all parts of my life, it gives me great pleasure to thank Kats, Cate, Robert and the rest of my (extended) family, but most of all, Tim.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

S.C.C. Stevens
December 12th, 2019

Publications

Some parts of this thesis appear in the following publications:

- [108] An Improved Point-Line Incidence Bound over Arbitrary Fields
  joint work with Frank de Zeeuw
  Bulletin of the London Mathematical Society 49.5 (2017), pp. 842-858

- [92] On the Energy Variant of the Sum-Product Conjecture
  joint work with Misha Rudnev and Ilya Shkredov
  To appear: Revista Matemática Iberoamericana, Electronically published on September 10, 2019

- [73] Bisector Energy and Pinned Distances in Positive Characteristic
  joint work with Brendan Murphy and Misha Rudnev
  arXiv preprint: 1908.04618

Sections containing reproduction are clearly indicated in the text.

Contents

List of Figures

List of Tables

Notation and conventions

Notation

Sets will always be finite and denoted by capital Latin letters. For example, $A = \{1, 2, 3\}$ denotes the set containing the numbers one, two and three. The cardinality of a set is denoted $|\cdot|$. e.g. $|A| = 3$.

$\mathbb{R}$, $\mathbb{C}$ etc. retain their usual meaning of the real and complex numbers. Similarly, $i = \sqrt{-1}$.

$\mathbb{F}$ denotes an arbitrary field. If the field is finite, then we write $\mathbb{F}_p$ to denote the prime residue field. The field $\mathbb{F}_q$ is a finite field of order q where q is a prime power. Throughout this work, p, q should be thought of as odd and large when referring to the characteristic of a field.

$\mathbb{F}^* = \mathbb{F} \setminus \{0\}$.
$\mathbb{F}^d$, for $d \geq 2$ an integer and $\mathbb{F}$ a field, denotes a d-dimensional vector space over $\mathbb{F}$. That is, $\mathbb{F}^d = \{(x_1, \ldots, x_d) \colon x_1, \ldots, x_d \in \mathbb{F}\}$.

For a set A we often write expressions of the form $A + A = \{a + b \colon a, b, \in A\}$. The meaning of this expression is typically intuitive, but will be defined throughout. However there are certain caveats, e.g. A/A is defined as the set $\{a/b \colon a, b \in A, b \neq 0\}$.

We write $\mathbb{1}_A(x) = 1$ if $x \in A$ and $\mathbb{1}_A(x) = 0$ if $x \notin A$. We also write

$\mathbb{1}_{f(x)=y} = 1$ if $f(x) = y$ and $\mathbb{1}_{f(x)=y} = 0$ if $f(x) \neq y$.

$r_{A+A}(x) := \sum_{a,b\in A} \mathbb{1}_{x=a+b}$ is the number of representations of an element x as $A + A$. The expression $r_{AA}(x)$ is defined similarly.

$\mathcal{P}$ typically denotes a set of points (in $\mathbb{F}^d$) and $\mathcal{L}$ typically denotes a set of lines. The set Π will typically refer to a set of planes. This notation will be always be restated.

The number of incidences between sets X and Y is

$$\mathcal{I}(X, Y) = |\{(x, y) \in X \times Y : x \cap y \neq \emptyset\}|..$$

The dot product between $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ is $x \cdot y = \sum_{i=1}^{d} x_i y_i = \langle x, y\rangle$.
The disjoint union is denoted $\sqcup$. The statement $A = B \sqcup C$ means that (i) $B \cap C = \emptyset$ and (ii) $A = B \cup C$.

Asymptotic Notation

When we are not concerned with constants, we use asymptotic notation as follows:

$$f(x) \ll g(x) \Leftrightarrow \exists\, C > 0, X \text{ such that } \forall\, x \geq X\,,\ \text{we have } f(x) \leq Cg(x)$$
$$f(x) \gg g(x) \Leftrightarrow \exists\, C > 0, X \text{ such that } \forall\, x \geq X\,,\ \text{we have } f(x) \geq Cg(x)\,.$$

Typically this notation will be with respect to the size of a set. E.g. the equation $|\{F(a, b) : a, b \in A\}| \gg |A|$ for some function F should be interpreted as asymptotic in the cardinality of the set A.

We write $f(x) \ll g(x)$ interchangeably as $f(x) = O(g(x))$, and similarly $f(x) \gg g(x)$ as $f(x) = \Omega(g(x))$.

This asymptotic notation hides constant terms, but often we will not be concerned with logarithmic terms:

$$f(x) \lesssim g(x) \Leftrightarrow \exists\, C_1 > 0, C_2 \in \mathbb{R} \text{ so that } \forall\, x \geq X\,,\ \text{we have } f(x) \leq C_1 \log^{C_2}(x) g(x)$$
$$f(x) \gtrsim g(x) \Leftrightarrow \exists\, C_1 > 0, C_2 \in \mathbb{R} \text{ so that } \forall\, x \geq X\,,\ \text{we have } f(x) \geq C_1 \log^{C_2}(x) g(x)\,.$$

The choice of logarithmic base does not matter in this notation since it change only the value of the implicit constants C_1 and C_2.

The motivation for hiding the logarithmic terms is that often we are so far from the true conjectured exponent of an expression that logarithmic losses are inconsequential in comparison.

1

Introduction

The unifying theme of this thesis is how incidence geometry can be used to approach questions in combinatorics. Typically the incidences in questions will be between points and lines in the plane, and the combinatorics in questions will be of an arithmetic flavour.

We are interested in asymptotic results: sets are *suitably large* to bypass exoticisms which occur at small scales. Since we are interested in the *arithmetic* structure, we require both an addition and multiplication operator, and so work in a field. If the field in which we work is finite, then we assume it is also suitably large. Geometric combinatorial questions in this area are of the type:

- How many incidences can there be between a set of points $\mathcal{P}$ and a set of lines $\mathcal{L}$ in terms of $|\mathcal{P}|$ and $|\mathcal{L}|$?
- How many incidences can there be between a set of points and lines if the points and/or lines have given structure?
- If a set of points and a set of lines has a maximal number of incidences, can we say anything about the structure of the configuration?
- Given a set of points in the plane, are there any structural features which are guaranteed to exist? We will ask how many distinct distances are there among all pairs of points in the set?

Over the real (and complex) numbers, the Szemerédi-Trotter theorem [110] provides an optimal bound on the number of incidences between points and lines. Over arbitrary fields $\mathbb{F}$ we still have points and lines, but are no longer

afforded the topology of the reals. When we ask questions over $\mathbb{F}$, we renounce the tools and methods developed over $\mathbb{R}$.

The combinatorial applications that we consider are varied, and it is often not obvious that incidences can be used in any way in their resolution. The applications we have in mind are of the following flavour:

- Given a set $A \subseteq \mathbb{F}$ for a field $\mathbb{F}$, can A simultaneously be 'additive' and 'multiplicative', for some suitable quantification of these terms? Generally one should have the motto in mind that addition and multiplication cannot coexist.

- What function $f : \mathbb{F}^d \to \mathbb{F}$ will satisfy $|\{f(a_1, \ldots, a_n) \colon a_i \in A\}| \gg |A|^\alpha$ for some $\alpha > 1$ and for any set $A \subseteq \mathbb{F}$? How big can α/d be?

In this thesis, we develop new incidence theorems over arbitrary fields $\mathbb{F}$. We investigate some applications, with an emphasis on the prime residue field $\mathbb{F}_p$. Since the cardinality of the set of interest will be large with respect to the characteristic of the field, we also require that $p \gg 1$. In particular, we examine the incompatibility of the coexistence of additive and multiplicative structure in a set. Finally, we find a new lower bound on the number of distinct distances a set of points in $\mathbb{F} \times \mathbb{F}$ must determine.

Structure of Thesis

- We begin with two introductory chapters: Chapter 2 offers an introduction to incidence bounds and Chapter 3 is a brief guide to arithmetic combinatorics and the sum-product problem. These two chapters contain both elementary and fundamental results.

- Chapter 4 is about incidence theorems in arbitrary fields. The key contribution of this chapter is a proof of two new incidence theorems. This is joint work with Frank de Zeeuw and the results are the content of the publication [108].

- Chapter 5 discusses applications of the incidence theorems proved in Chapter 4. The applications are contained in [108]; Chapter 5 provides additional explanation and context.

- In Chapter 6, we turn to additive combinatorics and an energy formulation of the sum-product problem. This work appears in the publication [92] and is joint with Misha Rudnev and Ilya Shkredov.

- Chapter 7 discusses two consequences of the results of Chapter 6: an application to expander functions, and a new sum-product result. Both of these results also appear in [92]. In this chapter we present a proof of the sum-product theorem avoiding auxiliary notation used elsewhere. It is hoped that this will aid the understanding and progress on this problem.

- Finally, Chapter 8 is about the pinned distance problem in arbitrary fields. This is joint work with Brendan Murphy and Misha Rudnev and is the content of the note [73].

2

A beginner's guide to incidence theorems

Throughout this thesis, numerous incidence bounds are used, in particular the Szemerédi-Trotter theorem and Rudnev's point-plane incidence bound. This chapter aims to serve as a reference point for these canonical results in the context of this work, and to describe their place within the mathematical literature. As such, details on the proofs of the results in this section are minimal. We refer the curious reader instead to the vast body of literature: e.g. [24, 98, 78, 65].

2.1 Incidences in two dimensions

Let us begin at the very beginning: The Elements of Euclid, Book I [18].

> I. A *point* is that which has no parts.
> II. A *line* is that without breadth.

Now, this definition is immediately understandable and fit for purpose over the reals, where the concept of width and breadth are inherited from our understanding of the physical world. However, over an arbitrary field, this intuition fails us, and we turn instead to another mathematical great. Descartes' introduction of Cartesian geometry allowed for an algebraic interpretation of points and lines, the definition of which extends over arbitrary fields. In modern language, a *point* is an element of a vector space of dimension $d \geq 1$

over a field $\mathbb{F}$: the notion of a point lacking parts is retained in this definition. Given two points $a, b \in \mathbb{F}^d$, a line is the set of elements of the form $ta + (1-t)b$ where $t \in \mathbb{F}$. Over the plane, a line can equivalently be defined algebraically: a line takes the form $\ell = \{(x, y) \in \mathbb{F}^2 : y = mx + c\}$ where $m, c \in \mathbb{F}$ are fixed and $m \neq 0$, or a line is vertical and takes the form $\ell = \{(v, y) : y \in \mathbb{F}\}$ for fixed $v \in \mathbb{F}$. When $\mathbb{F} = \mathbb{R}$, modern definitions and Euclid's classical definitions coincide. We refer to m in the definition of ℓ as the *gradient* of the line; the gradient of a vertical line is infinite.

A line is a collection of points; if we have a finite set of points $\mathcal{P}$ and a finite set of lines $\mathcal{L}$, we can ask how many of our points lie in our lines. If a point lies in a line, we say an *incidence* occurs and so we rephrase our combinatorial question by asking how many incidences there are between $\mathcal{P}$ and $\mathcal{L}$. We define the number of point-line incidences to be the quantity:

$$\mathcal{I}(\mathcal{P}, \mathcal{L}) : = \sum_{p \in \mathcal{P}} \sum_{\ell \in \mathcal{L}} \mathbb{1}_{p \in \ell} = \left| \{(p, \ell) \in \mathcal{P} \times \mathcal{L} : p \in \ell\} \right|.$$

For instance over $\mathbb{R}^2$, Figure 2.1 demonstrates an incidence count of two and Figure 2.2 demonstrates an incidence count of three. Unless explicitly stated otherwise, we shall restrict our investigation of incidences to the planar case.

Figure 2.1: Two incidences

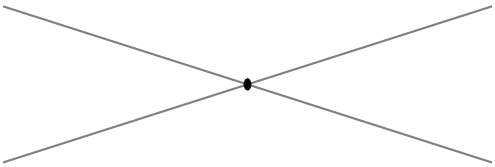
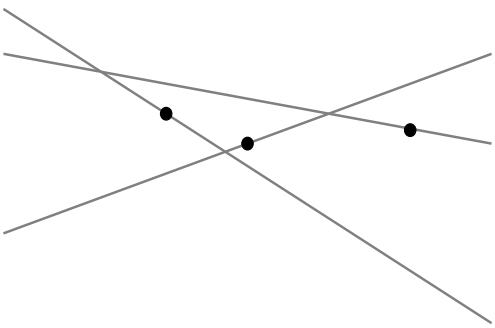


Figure 2.2: Three incidences



In $\mathbb{R}^2$ it is easy for finite sets of points and lines to have no incidences. In fact, this is the behaviour of 'typical sets'. However, if the sets of points and lines have many incidences, this is remarkable behaviour. Given that the sets of points and lines are both finite, it follows that the number of incidences between them is also finite. The first question we ask is therefore: how many incidences can there be?

First incidence bounds

It is clear that $\mathcal{I}(\mathcal{P}, \mathcal{L})$ is finite and of size at most $|\mathcal{P}||\mathcal{L}|$. This bound follows immediately from the fact that a line in $\mathcal{L}$ cannot contain more points than are present in the point set.

However, this bound can never[1] be realised because we have not accounted for the fact that two lines can intersect at only one point: if all points are collinear on $\ell \in \mathcal{L}$, contributing $|\mathcal{L}|$ incidences, then any other line in $\mathcal{L}$ can contribute at most one incidence to our count. This observation can be formalised and quantified via the Cauchy-Schwarz inequality to give the *trivial bound* on the number of incidences.

Lemma 2.1 (Trivial incidence bound)**.** *Let $\mathbb{F}$ be an arbitrary field. Let $\mathcal{P}$ and $\mathcal{L}$ be finite sets of $|\mathcal{P}|$ points and $|\mathcal{L}|$ lines in $\mathbb{F} \times \mathbb{F}$. Then*

$$\mathcal{I}(\mathcal{P}, \mathcal{L}) \leq \min\left(|\mathcal{L}|\sqrt{|\mathcal{P}|} + |\mathcal{P}|, |\mathcal{P}|\sqrt{|\mathcal{L}|} + |\mathcal{L}|\right).$$

Proof. From the Cauchy-Schwarz inequality we have:

$$\begin{aligned} I := \mathcal{I}(\mathcal{P}, \mathcal{L}) &\leq \sqrt{\sum_{p\in\mathcal{P}} 1 \cdot \sum_{p\in\mathcal{P}} \left(\sum_{\ell\in\mathcal{L}} \mathbb{1}_{p\in\ell}\right)^2} \\ &= |\mathcal{P}|^{1/2} \sqrt{\sum_{p\in\mathcal{P}} \sum_{\ell_1\in\mathcal{L}} \sum_{\ell_2\in\mathcal{L}} \mathbb{1}_{p\in\ell_1}\mathbb{1}_{p\in\ell_2}} \\ &= |\mathcal{P}|^{1/2} \sqrt{\sum_{p\in\mathcal{P}} \sum_{\ell_1\neq\ell_2\in\mathcal{L}} \mathbb{1}_{p\in\ell_1\cap\ell_2} + \sum_{p\in\mathcal{P}} \sum_{\ell\in\mathcal{L}} \mathbb{1}_{p\in\ell}} \\ &\leq |\mathcal{P}|^{1/2}\sqrt{|\mathcal{L}|(|\mathcal{L}|-1) + I} \\ &\leq |\mathcal{P}|^{1/2}\sqrt{|\mathcal{L}|^2 + I}\,, \end{aligned}$$

[1]Here, we use 'never' to mean never non-trivially; if $|\mathcal{P}| = 1$ then there are at most $|\mathcal{L}|$ incidences (by a pencil centred at the single point). By 'non-trivial' in this context we mean at least two points and two lines.

where the second multiplicative factor is simplified by the observation that two distinct lines can intersect at only one point.

Finally, we solve the ensuing quadratic inequality to obtain $2I \leq |\mathcal{P}| + \sqrt{|\mathcal{P}|^2 + 4|\mathcal{P}||\mathcal{L}|^2}$. Also note that $|\mathcal{P}| + \sqrt{|\mathcal{P}|^2 + 3|\mathcal{P}||\mathcal{L}|^2} \leq 2(|\mathcal{P}| + |\mathcal{L}|\sqrt{|\mathcal{P}|})$.

The second term in the statement of the lemma follows from the same calculation but with the roles of points and lines reversed. □

We note that because of the combinatorial nature of this proof, Lemma 2.1 holds in all fields.

This naturally leads to the following key question:

Question 2.2. *Is there a bound on $\mathcal{I}(\mathcal{P}, \mathcal{L})$ that is better than the trivial bound?*

The Szemerédi-Trotter Theorem

The question of the existence of an incidence bound better than the trivial bound was answered in the affirmative by Szemerédi and Trotter in 1983 [110] for the case $\mathbb{F} = \mathbb{R}$. It is difficult to overstate the importance of the following theorem and we will frequently refer to this theorem throughout this work.

Theorem 2.3 (Szemerédi-Trotter)**.** *Let $\mathcal{P} \subseteq \mathbb{R}^2$ be a finite set of $|\mathcal{P}|$ points and let $\mathcal{L}$ be a finite set of $|\mathcal{L}|$ lines in the real plane. Then there exists an absolute constant $c_{ST} > 0$ so that*

$$\mathcal{I}(\mathcal{P}, \mathcal{L}) \leq c_{ST} \left((|\mathcal{P}||\mathcal{L}|)^{2/3} + |\mathcal{P}| + |\mathcal{L}| \right) .$$

We will discuss the proof of Theorem 2.3 in Chapter 4, in particular, why various proofs of the Szemerédi-Trotter theorem do not extend to other fields. The current best value of c_{ST} remains an open problem; the state-of-the-art is $c_{ST} = 2.44\ldots$, a result of Ackerman [1].

Incidences between points and lines are present in the complex plane $\mathbb{C}^2$ and so we can also ask for a Szemerédi-Trotter theorem over $\mathbb{C}$. Since $\mathbb{R}$ is a subfield of $\mathbb{C}$, any example that demonstrates the sharpness of Theorem 2.3 also provides a lower bound for a complex Szemerédi-Trotter theorem; i.e. the best main term for an upper bound for the number of incidences between $|\mathcal{P}|$ points and $|\mathcal{L}|$ lines in $\mathbb{C}^2$ is at least $(|\mathcal{P}||\mathcal{L}|)^{2/3}$. Tóth [114] and Zahl [119] were successful in extending the Szemerédi-Trotter theorem to $\mathbb{C}^2$: they showed that, with possibly a different constant, Theorem 2.3 holds over $\mathbb{C}^2$. We refer to the

complex incidence theorem as the Szemerédi-Trotter theorem throughout this work.

The Szemerédi-Trotter Theorem has been applied in numerous other problems (see, e.g. the exposition [24]), and has been used as a vehicle to obtain incidences bounds between higher dimensional objects.

This theorem is optimal in the sense that there exist configurations of points and lines achieving each bound. We demonstrate this optimality with the following three examples in Figures 2.3–2.5:

Figure 2.3: The term $|\mathcal{P}|$ is attained: all $|\mathcal{P}|$ points are collinear on a line in $\mathcal{L}$

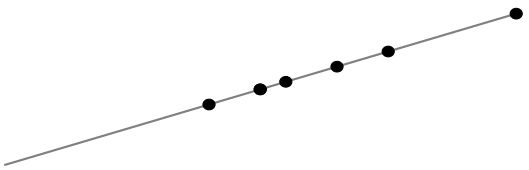


Figure 2.4: The term $|\mathcal{L}|$ is attained: all $|\mathcal{L}|$ lines are concurrent through a point in $\mathcal{P}$

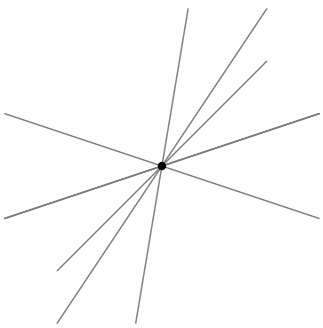
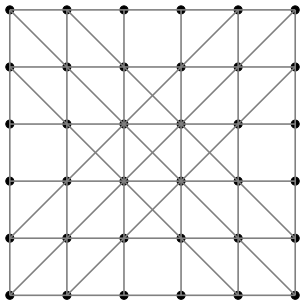


Figure 2.5: The term $(|\mathcal{P}||\mathcal{L}|)^{2/3}$ is attained: the $|\mathcal{P}|$ points are arranged in a grid on the integer lattice (if $|\mathcal{P}|$ is not a square, then the points are arranged to be 'as close to a square as possible'), and the $|\mathcal{L}|$ richest lines are chosen.



The number of k-rich points and lines

Definition 2.4. *Let $d \geq 2$ be an integer and let $\mathbb{F}$ be a field. Let $\mathcal{P} \subseteq \mathbb{F}^d$ be a set of points.*

A line ℓ in $\mathbb{F}^d$ is said to be k-rich with respect to $\mathcal{P}$ *if $|\ell \cap \mathcal{P}| \geq k$.*

We extend this definition to arbitrary sets and objects in $\mathbb{F}^d$. Let $S \subseteq \mathbb{F}^d$ be a set. Then $\theta \subseteq \mathbb{F}^d$ is k-rich with respect to S if $|\theta \cap S| \geq k$.

The Szemerédi-Trotter theorem gives an upper bound for the number of k-rich points with respect to a (finite) set of lines $\mathcal{L}$ lying in $\mathbb{C}^2$, and also an upper bound for the number of k-rich lines with respect to a finite set of points $\mathcal{P} \subseteq \mathbb{C}^2$.

Theorem 2.5 (k-rich points and lines)**.**

(i) Let $\mathcal{P} \subseteq \mathbb{R}^2$ be a finite set of points, and let $\mathcal{L}_k$ be the number of k-rich lines with respect to $\mathcal{P}$ for $k \geq 2$. Then

$$|\mathcal{L}_k| \ll \frac{|\mathcal{P}|^2}{k^3} + \frac{|\mathcal{P}|}{k}\,. \tag{2.1}$$

(ii) Let $\mathcal{L}$ be a finite set of real lines in $\mathbb{R}^2$, and let $\mathcal{P}_k$ be the number of k-rich points with respect to $\mathcal{L}$ for $k \geq 2$. Then

$$|\mathcal{P}_k| \ll \frac{|\mathcal{L}|^2}{k^3} + \frac{|\mathcal{L}|}{k}\,. \tag{2.2}$$

Proof. We prove only the estimate (2.1), and appeal to point-line duality as a proof of (2.2).

Let $\mathcal{P}, \mathcal{L}_k$ be as in the statement of the theorem. Without loss of generality, assume that $|\mathcal{P}| < (k/3)|\mathcal{L}_k|$ (else (2.2) is trivial). Then every line in $\mathcal{L}_k$ contains at least k points of $\mathcal{P}$. Hence $k|\mathcal{L}_k| \leq \mathcal{I}(\mathcal{P}, \mathcal{L}_k)$. We then apply the Szemerédi-Trotter theorem:

$$k|\mathcal{L}_k| \leq \mathcal{I}(\mathcal{P}, \mathcal{L}_k) \ll |\mathcal{P}|^{2/3}|\mathcal{L}_k|^{2/3} + |\mathcal{P}| + |\mathcal{L}_k|\,.$$

Since $k \geq 2$ and $|\mathcal{P}| \ll k|\mathcal{L}_k|$, we may disregard the final term at a cost of increasing the unspecified constant within the $\ll$ notation.

Rearranging this estimate then completes the proof of the theorem. □

Incidences over $\mathbb{R}^3$

Throughout this thesis, we will make references to a point-line incidence theorem in $\mathbb{R}^3$ of Guth and Katz [43]. This bound is particularly remarkable as it was used to prove an almost-tight bound for the Erdős distinct distance problem [29] in the plane, something which will be discussed in Chapter 8.

We will not use their theorem, but we state it here due it its importance in the literature. Guth and Katz proved the following sharp theorem.

Theorem 2.6 (Guth-Katz)**.** *Let $\mathcal{L}$ be a finite set of lines in $\mathbb{R}^3$, such that any doubly-ruled surface contains at most $O(|\mathcal{L}|^{1/2})$ lines in $\mathcal{L}$.*

Let $2 \leq k \leq |\mathcal{L}|^{1/2}$ and let $\mathcal{P}_k$ be the number of k-rich points with respect to $\mathcal{L}$. Then

$$|\mathcal{P}_k| \ll \frac{|\mathcal{L}|^{3/2}}{k^2}\,.$$

2.2 Incidences in arbitrary fields

How is $\mathbb{F}$ different to $\mathbb{R}$?

In arbitrary fields, the answer to Question 2.2 is in fact negative as can be seen by the following example. Take as the point set $\mathcal{P} = \mathbb{F}_p^2$ and choose the line set to be all possible $p^2 + p$ lines – that is, all p^2 lines of the form $y = mx + c$ as well as the p vertical lines $x = c$. Then every line contains p points and so there are $p^3 + p^2$ incidences, matching the trivial bound of Lemma 2.1.

At this point, one may wonder whether there is any hope of a positive answer to Question 2.2, by insisting that the point and line set avoid this pathological example. However, in extensions of prime fields, the situation is even worse, because the ground field now has finite subfields. Suppose the field $\mathbb{F}$ has a proper finite subfield $\mathbb{G}$; if we take $\mathcal{P} = \mathbb{G}^2$ and let $\mathcal{L}$ be the set of all lines of the form $ax + by + c = 0$ with coefficients $a, b, c \in \mathbb{G}$, then $\mathcal{I}(\mathcal{P}, \mathcal{L}) \approx |\mathbb{G}|^3$ but $|\mathcal{P}|^{2/3}|\mathcal{L}|^{2/3} \approx |\mathbb{G}|^{8/3}$.

From these two examples, it is clear that an incidence bound must take into account both the size of the field, and also the interaction between the point and line sets and any proper subfields. Accordingly, we rephrase Question 2.2 and ask instead: if we have a set of points $\mathcal{P}$ and a set of lines $\mathcal{L}$ so that the coordinates of $\mathcal{P}$ and the coefficients of equations defining lines in $\mathcal{L}$ do not lie in a proper subfield, is there an incidence bound better than the trivial incidence bound of Lemma 2.1?

Incidences over $\mathbb{F}^2$

When investigating incidence bounds over $\mathbb{F}^2$ for finite fields $\mathbb{F}$, there are three types of regime:

1. If both the sets of points and lines are large, then the plane is saturated and so we are forced to have incidences. Compare this with for example the real plane, where no matter how many points or lines we have, we can always place the points and lines such that they do not intersect.

2. In the special case $\mathbb{F} = \mathbb{F}_p$, if both the sets of points and lines are very small in cardinality with respect to p (that is, $|\mathcal{P}|, |\mathcal{L}| \ll \log\log\log(p)$), then a result of Grosu [40] tells us that '$\mathbb{F}_p$ is locally like $\mathbb{C}$': we can map $\mathbb{F}_p$ to $\mathbb{C}$ whilst preserving algebraic relations, and so in particular, we inherit the Szemerédi-Trotter theorem.

3. We have a significant number of points and lines, but they are not forced to have any interaction. This is the regime of interest throughout this thesis.

Since the results that we prove in Chapter 4 are relevant in the context of 'medium' sized sets of points and lines, we do not discuss them here.

Large sets of points

As the sizes of the point set and the line set grow, heuristically, the plane is 'saturated' by points and lines, forcing lots of incidences. In this regime, using a graph-theoretic framework and the eigenvalue-method, Vinh [116] was able to prove the following incidence bound:

Theorem 2.7 (Vinh [116])**.** *Let $\mathcal{P}$ be a finite set of points and $\mathcal{L}$ a finite set of lines in $\mathbb{F}_q^2$ for a prime power q. Then*

$$\mathcal{I}(\mathcal{P},\mathcal{L}) \leq \frac{|\mathcal{P}||\mathcal{L}|}{q} + (q|\mathcal{L}||\mathcal{P}|)^{1/2}\,.$$

This bound is able to perform best when $\max(|\mathcal{P}|,|\mathcal{L}|) > q$.

Small sets of points

A result of Grosu [40] of 2014 says that if a finite set of points $\mathcal{P} \subseteq \mathbb{F}_p \times \mathbb{F}_p$ and a finite set of lines $\mathcal{L} \subseteq \mathbb{F}_p \times \mathbb{F}_p$ are sufficiently small (specifically, $|\mathcal{P}|, |\mathcal{L}| < N$ and $5N < \log_2\log_6\log_{18}(p) - 1$), then we have a Szemerédi-Trotter type bound:

$$\mathcal{I}(\mathcal{P},\mathcal{L}) \leq CN^{4/3}\,.$$

As Grosu remarks in his paper, this bound, combined with [117, Theorem 2.3] of Vu, Wood and Wood, proves that the Szemerédi-Trotter bound is true over all integral domains of characteristic zero.

Incidences over $\mathbb{F}^3$

Points-Lines

When Guth and Katz published their ground-breaking result, one of the first questions was naturally to what extent their methods relied on the reals.

Kollár [57] was able to prove a weaker statement of Guth and Katz's result, [57, Theorem 2]. The methods of this weaker result, together with the results mentioned in [57, Paragraph 39] enabled Kollár to prove the following incidence theorem between points and lines in $\mathbb{F}^3$.

Theorem 2.8 (Points-Lines in $\mathbb{F}^3$)**.** *Let $\mathcal{L}$ be a set of distinct lines in $\mathbb{F}^3$ and $\mathcal{P}$ a set of distinct points in $\mathbb{F}^3$.*

If the characteristic of $\mathbb{F}$ is $p > 0$, assume that $p^3 > 6|\mathcal{P}|$. Let c be a constant so that no plane contains more than $O(\sqrt{|\mathcal{L}|})$ of the lines. Then the number of incidences satisfies

$$\mathcal{I}(\mathcal{P}, \mathcal{L}) \ll |\mathcal{L}||\mathcal{P}|^{1/3} + |\mathcal{P}|\,.$$

Points-Planes

For finite sets of points $\mathcal{P}$ and planes Π in $\mathbb{F}^3$ we define the number of incidences as

$$\mathcal{I}(\mathcal{P}, \Pi) = |\{(p, \pi) \in \mathcal{P} \times \Pi \colon p \in \pi\}|\,.$$

It is possible that all points in $\mathcal{P}$ are collinear on the line ℓ and that all planes in Π contain this line. In this situation, we have $\mathcal{I}(\mathcal{P}, \Pi) = |\mathcal{P}||\Pi|$; this number is maximal, and so we have that the trivial upper bound is realisable. However, in this example, the set of points is one-dimensional; if we insist that the points are instead truly three dimensional, then we are able to find a better incidence bound.

By interpreting 'truly three-dimensional' as a bound on the number of collinear points in $\mathcal{P}$, Rudnev [89] proved the following estimate:

Theorem 2.9 (Rudnev [89])**.** *Let $\mathcal{P}$ be a finite set of points in $\mathbb{F}^3$ and let Π be a finite set of planes in $\mathbb{F}^3$, with $|\mathcal{P}| \leq |\Pi|$. If $\mathbb{F}$ has positive characteristic p, suppose that $|\mathcal{P}| \ll p^2$. Let k be the maximum number of collinear points in $\mathcal{P}$. Then*

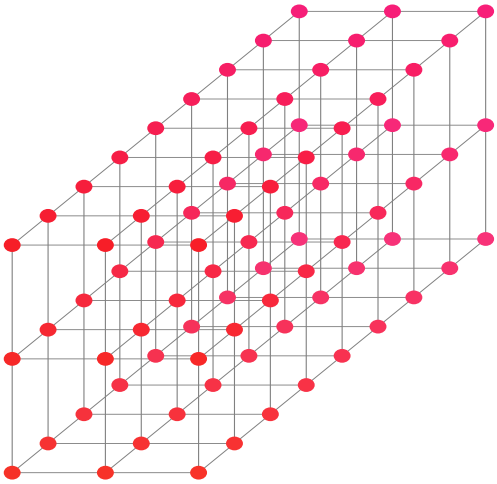
$$\mathcal{I}(\mathcal{P}, \Pi) \ll |\mathcal{P}|^{1/2}|\Pi| + k|\Pi|\,.$$

This theorem is sharp in the case of a cuboid grid, as the following example demonstrates.

Example 2.10. *Let $\mathcal{P} = \{1, \dots, N^2\} \times \{1, \dots, N\} \times \{1, \dots, N\} \subseteq \mathbb{R}^3$ be an integer lattice. We have $|\mathcal{P}| = N^4$.*

Let Π be the set of N^2 planes parallel to the yz-plane, and intersecting the x-axis at $x = 1, \dots, N^2$.

Figure 2.6: The point set $\mathcal{P} \subseteq \mathbb{R}^3$.



Each plane contains N^2 points of $\mathcal{P}$ and so $\mathcal{I}(\mathcal{P}, \Pi) = N^4$. Moreover, there are at most N^2 collinear lines in $\mathcal{P}$, and so the number of incidences meets the estimate of Rudnev's bound.

The proof of Rudnev's theorem is based on the ground-breaking result of Guth and Katz [43]. Rudnev maps both points and planes to two families of planes in the Klein quadric (a four-dimensional variety in $\mathbb{PF}^5$): an incidence between a point p and a plane q in $\mathbb{F}^3$ becomes a line defined by the intersection of the two planes π_p,π_q in the Klein quadric, where π_p and π_q are the images of the point $p \in \mathbb{F}^3$ and the plane $\pi \in \mathbb{F}^3$ respectively. The ensuing number of lines is estimated using the algebraic machinery of Guth and Katz, as stated in Kollar's theorem (Theorem 2.8).

A recent work of de Zeeuw [120] simplified Rudnev's proof, by mapping intersections between points and planes directly to intersections between lines in $\mathbb{F}^3$.

We can equivalently restate this theorem with the condition $|\mathcal{P}| \leq p^2$; changing from $\ll$ to $\leq$ changes only the constant in (2.9). We also have the dual form of this theorem, where the roles of points and planes are interchanged. The condition at most k collinear points is replaced with the requirement that at most k planes intersect in the same line.

Theorem 2.9 has yielded numerous applications, as is exposited in a paper of Rudnev [90] and references therein.

Over $\mathbb{R}$, better bounds are known for small values of k: e.g. Edelsbrunner et al. [20] proved a tight bound if no three points are collinear; Elekes and

Toth [28] have a (different) tight bound under the assumption that no line in a plane supports more than a constant proportion of incidences in that plane. Basit and Sheffer [6] hold the best result in this direction: for any finite sets $\mathcal{P}, \Pi \subseteq \mathbb{R}^3$ of points and planes respectively, the number of incidences between $\mathcal{P}$ and Π is at most $O(|\mathcal{P}|^{4/5+\epsilon}|\Pi|^{3/5}k^{2/5} + |\Pi| + |\mathcal{P}|k)$, where k is the maximum number of concurrent planes in Π.

2.3 Collinear triples

A collinear triple is a triple of points lying on the same line. To be precise, we make this definition with regards to points over an arbitrary field, excluding the trivial collinear triples of the form $(a, a, a) \in (\mathbb{F}^2)^3$ and triples like $(a, a, b) \in (\mathbb{F}^2)^3$ etc. in our count.

Definition 2.11. *A* collinear triple *is an ordered triple of distinct points (a, b, c) such that the (unique) line passing through a and b contains c.*

For a point set $\mathcal{P} \subseteq \mathbb{F}^2$, we define the number of collinear triples to be:

$$T(\mathcal{P}) := |\{(a, b, c) \in \mathcal{P}^3 : \text{ there is a line } \ell \subseteq \mathbb{F}^2 \text{ with } a, b, c \in \ell\}| . \tag{2.3}$$

We have the immediate bound $T(\mathcal{P}) \leq \binom{|\mathcal{P}|}{3}$.

We will be particularly interested in the count of collinear triples determined by a finite set of points with a Cartesian product structure. Note that for $A, B \subseteq \mathbb{F}$, $T(A \times B) = T(B \times A)$.

In a Cartesian product structure, we automatically inherit the horizontal and vertical collinear triples; these unavoidable terms will dominate the count of collinear triples unless $|A|, |B|$ are of comparable cardinalities. Hence we suppose that there exist constants c, C so that $cN \leq |A|, |B| \leq CN$.

The number of collinear triples of a Cartesian product is a quantity of interest because we have an optimal bound in this setting. This is an elementary consequence of the Szemerédi-Trotter theorem, first observed by Elekes and Ruzsa [26].

Theorem 2.12 (Elekes-Ruzsa [26])**.** *Let $A, B \subseteq \mathbb{R}$ be finite sets so that $|A|, |B| \sim N$. Then*

$$T(A \times B) \lesssim N^4 .$$

Proof sketch. Let $\mathcal{L}_k$ be the set of lines containing between 2^k and 2^{k+1} points of $A \times B$ for $1 \leq k \leq \lceil \log_2(N) \rceil$. Using the k-rich variant of the Szemerédi-Trotter theorem, Theorem 4.8, we have:

$$\begin{aligned} T(A \times B) &= \sum_{k=1}^{\lceil \log_2(|B|) \rceil} \binom{2^{k+1}}{3} |\mathcal{L}_k| \leq \sum_k \binom{2^{k+1}}{3} \frac{|A|^2|B|^2}{(2^k)^3} + \binom{2^{k+1}}{3} \frac{|A||B|}{2^k} \\ &\ll \sum_{k=1}^{\log(|B|)} |A|^2|B|^2 + |A||B|2^{2k} \,. \end{aligned}$$

□

When $\mathcal{P} = A \times B$, and $cN \leq |A|, |B| \leq CN$, we have yet another characterisation of the number of collinear triples.

We introduce the *representation function* of an element $r_X(x)$ to denote the number of representations x has as an element of X. For example, for a finite set A, $r_{A+A}(x) := \{(a,b) \in A^2 \colon x = a + b\}$. In this section we are interested in representations of elements of the set $\frac{B-b}{A-a}$ for sets A, B and fixed elements a, b. We define

$$r_{\frac{B-b}{A-a}}(x) := |\{(\alpha, \beta) \in A \times B \colon x = \frac{\beta - b}{\alpha - a}\}| \,.$$

Let $T' = T'(A \times B)$ be

$$T' = \sum_{x \in \mathbb{F}} \sum_{(a,b) \in A \times B} r^2_{\frac{B-b}{A-a}}(x) \,.$$

The quantities T and T' are related by $T(A \times B) \sim T'(A \times B) + N^4$. Indeed

$$\begin{aligned} T'(A \times B) &= \sum_{m \in \mathbb{F}} \sum_{(c_1,c_2) \in A \times B} r^2_{\frac{c_2 - B}{c_1 - A}}(m) \\ &= \sum_{(c_1,c_2) \in A \times B} \sum_{m \in \mathbb{F}} \left(\sum_{\substack{(a_1,a_2) \in A \times B \\ a_1 \neq c_1}} \mathbb{1}_{m = \frac{c_2 - a_2}{c_1 - a_1}} \right)^2 \\ &= \sum_{\lambda, m \in \mathbb{F}} \sum_{(c_1,c_2) \in A \times B} \mathbb{1}_{c_2 = mc_1 + \lambda} \left(\sum_{\substack{(a_1,a_2) \in A \times B \\ a_1 \neq c_1}} \mathbb{1}_{a_2 = ma_1 + \lambda} \right)^2 \\ &= |\{(a,b,c) \in (A \times B)^4 : a_1, b_1 \neq c_1, a, b, c \in \ell_{m,\lambda}\}| \,, \end{aligned}$$

where $\ell_{m,\lambda}$ is the line defined by $y = mx + \lambda$.

In the count of $T(A\times B)$ we must both include the $|A|\binom{|B|}{3}$ vertical collinear triples omitted from the count of $T'(A\times B)$, and also exclude the $O(N^4)$ 'trivial' collinear triples: we have added the triple (c,a,a) to the count of $T'(A\times B)$.

Hence, we (somewhat inaccurately) consider $T(A,B)$ and $T'(A,B)$ to be equivalent because Theorem 2.12 provides the same asymptotic bounds for both quantities: $N^4 \ll T(A,B), T'(A,B) \lesssim N^4$ where $|A|, |B| \sim N$.

2.4 Addendum: Projective transformations

Thus far, our definition of the two dimensional vector space $\mathbb{F}^2$, allows for parallel lines, that is, distinct lines sharing the same gradient. In the *projective plane* $\mathbb{PF}^2$, this possibility is excluded by the introduction of an additional line and additional points 'at infinity'.

An abstract projective plane is defined to be a set of points $\mathcal{P}$ and a set of lines $\mathcal{L}$ together with an incidence relation satisfying: (i) any two lines are incident to exactly one point; (ii) any two points are incident to exactly one line; (iii) there exist four points of $\mathcal{P}$ such that any line of $\mathcal{L}$ is incident with at most two of them.

To define a projective plane over $\mathbb{F}$, we start with the affine plane over $\mathbb{F}$ (with points $(a,b)\in\mathbb{F}^2$), and we add a 'point at infinity' to every collection of parallel lines, ensuring that each 'point at infinity' lies on a 'line at infinity'. This generates the projective plane $\mathbb{PF}^2$.

As coordinates, points in $\mathbb{PF}^2$ are defined via an equivalence relation $\sim$ where $[x:y:z]\sim[x':y':z']$ if there exists $k\neq 0$ such that $(x,y,z)=(kx,ky,kz)$. Then points in $\mathbb{PF}^2$ are

$$\{[x:y:z]/\sim : (x,y,z)\in\mathbb{F}^3\setminus\{0,0,0\}\}.$$

Given a triple $(a,b,c)\in\mathbb{F}^3\setminus\{(0,0,0)\}$, we define the line $\ell_{a,b,c}$ as

$$\ell_{a,b,c}=\{[x:y:z]/\sim : (x,y,z)\in\mathbb{F}^3\setminus\{(0,0,0\} : ax+by+cz=0\}$$

The set of lines in $\mathbb{PF}^2$ is $\{\ell_{a,b,c} : [a:b:c] \text{ is a point in } \mathbb{PF}^3\}$.

The point $[x:y:Z]$ is incident to the line $\ell_{a,b,c}$ if $[x:y:Z]\in\ell_{a,b,c}$.

We embed the affine plane $\mathbb{F}^2$ into $\mathbb{PF}^2$ by mapping the point $(x,y)\in\mathbb{F}^2\mapsto[x:y:1]\in\mathbb{PF}^2$. The line in $\mathbb{F}^2$ defined as the collection of points (x,y) satisfying the linear relation $ax+by+c=0$ is mapped to the line in

$\mathbb{PF}^2$ defined as the collection of points $[x : y : Z]$ satisfying the linear relation $ax + by + cz = 0$.

The points at infinity are of the form $[x : y : 0]$ where $(x, y) \neq (0, 0)$; these are orthogonal to the vector $(0, 0, c)$ for any $c \neq 0$ and form a line *at infinity*, denoted ℓ_∞.

Embedding a configuration in $\mathbb{F}^2$ into $\mathbb{PF}^2$ preserves the incidence structure of the configuration – if a point lies in a line in $\mathbb{F}^2$, it does so in $\mathbb{PF}^2$ – and so we will often turn to projective geometry when resolving incidence questions.

If we return to the abstract definition of a projective plane, it is clear that the roles of points and lines can be reversed in (i) and (ii). The third condition, that there are at least four non-collinear points so that any line is incident to at most two points, can be equivalently rephrased as: (iii)' there exist at least four non-concurrent lines so that any point is incident to at most two lines. This motivates the concept of *point-line duality* in the projective plane: in any statement involving points and lines in a projective plane, the roles of points and lines can be reversed.

For example, the projective point $[a : b : c] \in \mathbb{PF}^2$ is dual to the projective line $\ell_{a,b,c} := \{[x : y : z] \in \mathbb{PF}^2 \colon ax + by + cz = 0\}$. From this, it follows that collinear points are dual to concurrent lines.

Projective transformations will be used in Chapter 4. We use the book of Richter-Gebert [85] which gives a more thorough and leisurely introduction to the projective plane and projective transformations, and for the following facts.

Definition 2.13. *A projective transformation is a bijective linear map* $\mathbb{PF}^2 \to \mathbb{PF}^2$.

We call two sets projectively equivalent if there is a projective transformation that maps one bijectively to the other.

Importantly, a projective transformation maps lines to lines and hence preserves collinearity of points.

There are two points α, β on the line at infinity such that all lines through α (except for ℓ_∞) are horizontal lines in the affine plane, and the lines through β (except for ℓ_∞) are vertical lines in the affine plane.

Finally, for any two points $p, q \in \mathbb{PF}^2$ there is a projective transformation that sends p and q to α and β (see for instance [85, Theorem 3.4]).

3

A beginner's guide to arithmetic structure

3.1 Arithmetic Structure in a Set and the Sum-Product Conjecture

The main object of interest throughout this work will be a finite set A that is a subset of a field $\mathbb{F}$. We are interested in the asymptotic behaviour of the set in terms of its cardinality $|A|$. The reader is encouraged to have a concrete favourite field in mind, although should be warned that this section has been written with the intention of applications to the author's preferred fields, namely the reals $\mathbb{R}$ and prime residue fields $\mathbb{F}_p$. In the specific context of $\mathbb{F}_p$, note that since $|A|$ is asymptotically large, so too is p.

However, we begin with a simpler object, namely a finite set $A \subset \mathbb{G}$, where $(\mathbb{G}, +)$ is a group. Here, $\mathbb{G}$ is an additive group and so is endowed with additive structure, namely that $\mathbb{G}$ is closed under addition. If A is also closed under addition, then it is a subgroup (since it is finite) and so it is clear that it is an object of interest. But if A is 'nearly closed' under addition, then intuitively we might think that it is 'subgroup-like', and endowed with structure, where the quantification of 'additive structure' correlates with the proximity of A to an additively closed set. When $\mathbb{G} = \mathbb{Z}$, this intuition is formalised by Freiman's theorem [35], which states that if A is almost closed with respect to addition, then in fact A is contained inside a generalised

arithmetic progression[1] of controlled size and dimension. If $\mathbb{G}$ is an abelian group, Green and Ruzsa [39] prove a (quantified version of the) statement that says if A is almost closed with respect to addition, then A is contained in an object with a very particular structure: $A \subseteq H + P$ where H is a subgroup and P is a generalised arithmetic progression of controlled size and dimension. The literature is rich with quantifications of this type of statement, see e.g. [75]. The main interpretation of this type of result should be that if $A \subseteq \mathbb{G}$ exhibits additive structure (i.e. A is almost closed with respect to addition), then A is a large subset of something very structured.

In this group setting, there is only one type of structure that the set A can exhibit. When we turn to the case of a set in a field $\mathbb{F}$, then A, the set in question, can exhibit both additive and multiplicative structure. As we will be interested in asymptotic results, we can make the simplifying assumption that $0 \notin A$ and so results of the type mentioned in the previous paragraph hold with the multiplicative group $\mathbb{G} = \mathbb{F}^*$. Consequently, if A is additively closed, it is contained in a structured 'additive object', and if it is multiplicatively closed, then it is contained in a structured 'multiplicative object'. Of course, A could be closed with respect to both addition and multiplicative, in the case where A is a sub-field of $\mathbb{F}$. If we rule out this exceptional example, the question of what structure A can have if it is closed with respect to both operations should be preceded with a more fundamental query: is it *possible* for A to be almost closed with respect to both addition and multiplication if A is not a sub-field?

Quantifying Structure: Cardinality

In this subsection, we formalise the question of whether addition and multiplication can coexist in a set.

Given a finite set of elements $A \subseteq \mathbb{F}$, we define the *sum set* $A + A$ to be the set of all pairwise sums of elements of A; the *product set* AA is similarly created via pairwise products:

$$A + A := \{a + b : a, b \in A\} \quad \text{and} \quad AA := \{ab : a, b \in A\}.$$

In this notation, A is an additive (respectively multiplicative) subgroup if $A + A = A$ (resp. if $AA = A$). It is easy to see that this implies that

[1] A generalised arithmetic progression of dimension d and volume $\prod_{i=1}^{d} N_i$ is a set of the form

$$P = \{a + n_1 v_1 + \cdots + n_d v_d : 0 \leq n_i \leq N_i \text{for all } 1 \leq i \leq d\}$$

where $a, n_1, \ldots, n_d, v_1, \ldots, v_d$ are elements of the ambient space.

$|A+A| = |A|$; in fact, these statements are almost equivalent: if $|A+A| = |A|$ then A is a coset of of subgroup – see [111, Proposition 2.2]. If $|A+A|/|A|$ is small, then we think of A being *almost* additively closed. Hence the question of whether A can be simultaneously additively and multiplicatively structured then becomes the question of whether $|A+A|/|A|$ and $|AA|/|A|$ can both be simultaneously small.

We have the combinatorial bounds

$$|A| \leq |A+A|, |AA| \leq \binom{|A|+1}{2}.$$

To understand how the cardinalities of the sum and product set interact, we consider three examples of sets in $\mathbb{Z}$ of cardinality $N \gg 1$.

Example 3.1. [Random Set]

Let T be an integer so that $N^4/T \to 0$ as $N \to \infty$. We will construct a random set $R \subseteq \{1, \dots, T\}$ so that $\mathbb{P}(t \in R) = NT^{-1}$ for each $t = 1, \dots, T$.

Note that $\mathbb{E}(|R|) = N$ and $\mathbb{E}(|R+R|) = \mathbb{R}(|RR|) = \binom{N+1}{2}$: we do not expect any element of $R+R$ or RR to have multiple representations as sums from R.

A probabilistic argument which we leave to the reader formalises this example – see e.g. [111, Chapters 1 and 2].

This example shows that it is easy for both the $|A+A|/|A|$ and $|AA|/|A|$ to be *maximal*. We consider now sets in which one of these quantities is minimal.

Example 3.2. [Arithmetic Progression]

Let $A = \{1, 2, \dots, N\} \subseteq \mathbb{N}$.

The sum set is $A+A = \{2, 3, \dots, 2N\}$, of size $2N-1$. We note that this bound is an extremal lower bound for the sum-set.

The product set AA is contained in $\{1, \dots, N^2\}$. It is not hard to show that $|AA| \gg (|A|/\log(|A|))^2$ (see e.g. [76, Lemma 5]). Estimating $|AA|$ precisely is known as the multiplication table problem and the best results in this area are due to Ford [34] who proves that $|AA| \sim \frac{|A|^2}{(\log|A|)^\delta (\log\log|A|)^{3/2}}$ where $\delta = 1 - (1 + \log\log(2))/\log(2) = 0.086071\dots$.

In Example 3.2, the quantity $|A+A|/|A|$ attains its minimum, but its multiplicative analogue is maximal (up to logarithmic factors).

Example 3.3. [Geometric Progression]

Let $B = \{1, 2, 2^2, \dots, 2^N\}$.

Notice that $\log(B) := \{\log(b) : b \in B\}$ *is the set* A *of Example 3.2. Hence the product set satisfies* $BB = 2^{\log(B)+\log(B)} = 2^{(A+A)} = \{2^2, 2^3, \dots, 2^{2N}\}$.

The sum set $B + B$ *is contained in* $\{2, \dots, 2^{2N}\}$. *All elements are distinct, which can be seen by turning to the binary representations of the elements. Hence* $|B + B| = \binom{|B|+1}{2}$.

In the example of a random set, there is intuitively 'no structure', whereas it is clear that the arithmetic progression is an additively structured object, and the geometric progression is a multiplicative structured set.

In Examples 3.1 to 3.3, at least one of the sum-set or product-set is large. Erdős and Szemerédi [31] conjectured that this behaviour is typical of all sets in $\mathbb{N}$. Their conjecture has since been extended to all finite sets in $\mathbb{R}$ and it is this version to which we refer as the *sum-product problem*.

Conjecture 3.4 (Sum-Product Problem)**.** *For all* $\epsilon > 0$*, there exists a constant* $C = C(\epsilon) > 0$ *so that for all finite sets* $A \subseteq \mathbb{R}$

$$\max(|A + A|, |AA|) \geq C|A|^{2-\epsilon}.$$

The sum-product problem is further addressed in Chapters 5 and 7 and is a major motivating question of this thesis, in particular the extent to which a sum-product phenomenon can occur in arbitrary fields.

Progress on the Sum-Product Phenomenon

Erdős and Szemerédi [31] proved a modicum towards Conjecture 3.4 (they showed a qualitative statement towards the conjecture), and subsequent work of Nathanson [74] and Ford [33] respectively quantified and finessed their approach. This problem has seen two milestone results, both of which changed the approach to this problem, each remarkable in their elegance and simplicity. The first, by Elekes [25], connected the sum-product problem with the Szemerédi-Trotter theorem and thus incidence geometry. (This approach was later developed further by Solymosi [105].) The second development, by Solymosi [104], used elementary geometry to bound the multiplicative energy of a set by its sumset. This will be discussed in Section 3.3. This approach was further developed by Konyagin and Shkredov [60, 58], and then by Rudnev, Shkredov and the author [92]. This improvement is described in Chapter 5.

The best exponent to date is by Shakan [97], giving a lower bound of $|A|^{4/3+c}$, where $c = 5/5277$.

In this chapter, we state and sketch the results of Elekes [28] and Solymosi [104]. The latter will be stated after developing the necessary terminology.

Theorem 3.5 (Elekes [28])**.** *Let $A \subseteq \mathbb{R}$ be a finite set. Then*

$$|A + A|^2|AA|^2 \gg |A|^5 \tag{3.1}$$

and in particular

$$\max(|A + A|, |AA|) \gg |A|^{5/4} \tag{3.2}$$

Proof. Let $\mathcal{P} = (A + A) \times AA \subseteq \mathbb{R}^2$ be a set of points, and let $\mathcal{L} = \{\ell_{ab} \colon a, b \in A\}$ be a set of lines where

$$\ell_{ab} := \{(x, y) \in \mathbb{R}^2 : y = a(x - b)\}.$$

A line $\ell_{ab} \in \mathcal{L}$ contains the point $(c + b, ac) \in \mathcal{P}$ for every $c \in A$; ℓ_{ab} contains $|A|$ points of $\mathcal{P}$. Hence, using the Szemerédi-Trotter theorem:

$$|A||\mathcal{L}| \leq \mathcal{I}(\mathcal{P}, \mathcal{L}) \ll (|\mathcal{P}||\mathcal{L}|)^{2/3} + |\mathcal{P}| + |\mathcal{L}|.$$

The proof then follows since $|\mathcal{P}| = |A + A||AA|$ and $|\mathcal{L}| = |A|^2$. □

Addendum: The Plünnecke-Ruzsa Inequality

Having defined the sum and product sets of A, it is natural to ask about, say, $A + A + A$ etc. When introducing the sum-product problem in [31], Erdős and Szemerédi conjectured more than the statement of Conjecture 3.4: at least in the integers, either the *k-fold* sum-set $A + A + \cdots + A$ or the k-fold product set $AA \cdots A$ should be of cardinality $|A|^{k-\epsilon}$. We do not discuss this stronger version in this work, but mention a cornerstone of additive combinatorics which will hopefully convince the reader that the sum-product conjecture captures any structure appearing in k-fold sum and product sets: the Plünnecke-Ruzsa theorem.

We define the k-fold sum-set of A to be

$$kA := \{a_1 + \cdots + a_k \colon a_1, \ldots, a_k \in A\}.$$

Theorem 3.6 (Plünnecke-Ruzsa [84, 93])**.** *Suppose A is a subset of an abelian group $\mathbb{G}$. If $|A + A| \leq K|A|$ then for any integers $m, n \geq 0$, $|mA - nA| \leq K^{m+n}|A|$.*

An interpretation of the Plünnecke-Ruzsa theorem is that the 2-fold sum-set of A is a suitable object to study to determine whether A has additive structure. An analogous definition, statement and interpretation can be made for product sets of A. The Plünnecke-Ruzsa theorem was originally proved by Plünnecke [84] and rediscovered by Ruzsa [93]. Both proofs use a graph theoretic technique; in 2012 Petridis [80] discovered an elementary approach to this important theorem.

3.2 Arithmetic Energy

The sum-product phenomenon measures to what extent a set can be both additive and multiplicative. However, studying the cardinality of the sum-set and product-set is not the sole way to extract this structure.

Definitions and Examples

To gain some intuition, consider the sum-set of A. To create the sum-set, we can create at most $\binom{|A|}{2} + |A|$ different sums; if the sum-set is small, then many of these sums must coincide, namely there must be many quadruples $(a, b, c, d) \in A^4$ so that $a+b = c+d$. The number of these quadruples is termed the (additive) energy of A (a term coined by Tao and Vu [111]).

Definition 3.7. *The additive energy between sets A and B is*

$$\mathsf{E}^+(A, B) = |\{(a_1, a_2, b_1, b_2) \in A^2 \times B^2 : a_1 + b_1 = a_2 + b_2\}|.$$

The multiplicative energy is

$$\mathsf{E}^\times(A, B) = |\{(a_1, a_2, b_1, b_2) \in A^2 \times B^2 : a_1 b_1 = a_2 b_2\}|\,.$$

We write $\mathsf{E}^+(A, A) = \mathsf{E}^+(A)$; if this quantity is large, then A has an additive structure, and if it is small then A has little additive structure; the same principle holds regarding $\mathsf{E}^\times(A)$ and multiplicative structure. We can replace addition with subtraction in the definition of the additive energy; if we assume that $0 \notin A$, then we may replace multiplication with division in the definition of multiplicative energy.

We have the bounds

$$|A|^2 \leq \mathsf{E}^+(A), \mathsf{E}^\times(A) \leq |A|^3\,.$$

These bounds follow from two observations: firstly, there are $|A|^2$ quadruples of the form $(a, b, a, b) \in A^4$ contributing to the energy count and secondly, fixing three of the elements in a quadruple of the support of the energy determines the fourth.

A motto to keep in mind should be *a set with large energy has structure.* Knowing that the additive (resp. multiplicative) energy of a set A is small enables us to deduce that the sum-set (resp. product-set) has large cardinality in terms of $|A|$. However, the converse does not hold, as the following examples demonstrate.

Example 3.8 (Small sum-set, large additive energy)**.** *Consider* $A = \{1, \dots, n\}$. $A + A = \{2, \dots, 2n-1\}$ *and so* $|A + A| \ll |A|$.
The energy of A *is maximal:* $\mathsf{E}^+(A) \gg |A|^3$.

Example 3.9 (Large sum-set, small additive energy)**.** *Consider* $A = \{2, 2^2, \dots, 2^n\}$, *a geometric progression. The sum-set of this is* $|A + A| \gg |A|^2$ *(see e.g. [111, Chapter 2], and* $\mathsf{E}^+(A) \ll |A|^2$ *(a consequence of the uniqueness of binary representations of integers).*

Example 3.10 (Large sum-set, large additive energy)**.** *Consider*

$$A = \{1, \dots, n\} \cup \{2^n, 2^{n+1}, \dots, 2^{2n}\},$$

the union of an additive progression and a geometric progression.
Then $|A + A| \geq |\{2^n, 2^{n+1}, \dots, 2^{2n}\} + \{2^n, 2^{n+1}, \dots, 2^{2n}\}| \gtrsim |A|^2$ *(the sum-set is large). Both the additive and multiplicative energies of this set are large:* $\mathsf{E}^+(A), \mathsf{E}^\times(A) \gg |A|^3$

In Example 3.10, we have a set which clearly contains both additive and multiplicative structure, which we capture by having large multiplicative and additive energy. However, both the sum-set and product-set are large, indicating that the cardinality of sets is not a sufficient measure of the structure of a set. This example shows that the arithmetic energy of a set is a finer measure of arithmetic structure.

Higher energies

We have defined the cardinality of a set as a first moment quantification and we have defined the additive energy of A as a second moment measure:

$$|A|^2 = \sum_{x \in A-A} r_{A-A}(x); \qquad \mathsf{E}^+(A) = \sum_{x \in A-A} r^2_{A-A}(x).$$

Naturally we can continue in this fashion, and define the *higher energies* of a set for $k \geq 0$:

$$\mathsf{E}_k^+(A) := \sum_{x \in A-A} r_{A-A}^k(x); \qquad \mathsf{E}_k^\times(A) := \sum_{x \in A/A} r_{A/A}^k(x)\,.$$

Higher energies were first thoroughly treated in a work of Schoen and Shkredov [96].

Geometric interpretation of energy

The quantities $\mathsf{E}^+(A)$ and $\mathsf{E}^\times(A)$ can be interpreted geometrically. This will be useful in the sequel.

Consider first the multiplicative energy:

$$\mathsf{E}^\times(A,B) = \sum_{\lambda \in A/B} |\{(a,a',b,b') \in A^2 \times B^2 : a/b = a'/b' = \lambda\}|\,.$$

If we think of λ in this expression as referring to the line ℓ_λ defined by $y = \lambda x$, then to calculate the multiplicative energy, sum for all $\lambda \in A/B$ the number of pairs of points of $A \times B$ lying on the line ℓ_λ. Note that we consider the 'trivial pair' $((a, \lambda a), (a\lambda a)$ in this count.

The count for the additive energy is the similar, except this time we sum either over parallel lines of the form $y = c + x$ where $c \in A + B$, corresponding to the definition $\mathsf{E}^+(A,B) = \sum_{x \in A+B} r_{A+B}^2(x)$; or else we sum over parallel lines of the form $y = c - x$ if $c \in A - B$, corresponding to the equivalent definition $\mathsf{E}^+(A,B) = \sum_{x \in A-B} r_{A+B}^2(x)$.

Note that a dual geometric interpretation could be made for the product set (in place of the ratio set A/B) $\mathsf{E}^\times(A,B) = \sum_{p \in AB} r_{AB}^2(x)$. However this is more complicated than the equivalent ratio-set interpretation of the energy.

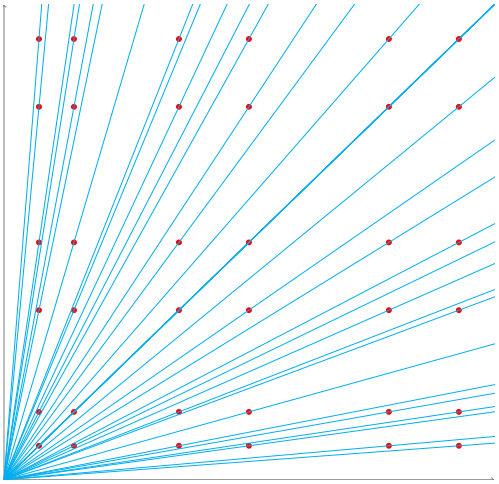


Figure 3.1: Geometric interpretation of multiplicative energy

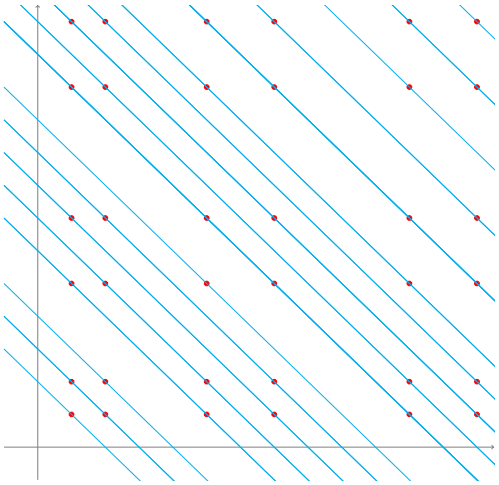


Figure 3.2: Geometric interpretation of additive energy

For $\mathsf{E}_2^+(A)$, the roles of addition and subtraction are interchangeable; for higher energies this is no longer the case. We could (but will not) similarly define the higher energy with respect to the functions r_{A+A} or r_{AA} etc. If the operative superscript is omitted for $A \subseteq \mathbb{F}$, then the energy in question will always be the additive energy $\mathsf{E}_k(A) = \mathsf{E}_k^+(A)$.

The k-th energy is not limited to integer values of k. However, when $k \in \mathbb{N}$, then we can benefit from a geometric interpretation of the k-th energy. Recall that the ordinary $\mathsf{E}_2(A)$ energy could be found by summing, over all appropriate slopes (where these slopes are either 'multiplicative' and all pass

through the origin, or slopes are all 'additive' and are parallel with slope -1), the number of pairs of points of $A \times A$ per slope.

This count extends to k-th energy, after the amendment that we sum over the appropriate slopes the number of k-tuples of points of $A \times A$ lying on each slope.

In the case when $k \in \mathbb{N}$, we have the alternative definition:

$$\mathsf{E}_k^+(A) := |\{(a_1, b_1, \ldots, a_k, b_k) \in A^{2k} \colon a_1 - b_1 = \cdots = a_k - b_k\}\,.$$

We have the immediate bounds:

$$|A|^k \leq \mathsf{E}_k(A) \leq |A|^{k+1}\,.$$

The role of the third higher energy is particularly useful in the reals. This is because the quantity $\mathsf{E}_3(A)$ can be associated to the count of collinear triples, and the Szemerédi-Trotter theorem provides an optimal upper bound on this count. We do not prove this here, but this notion is captured in Chapter 7 in Theorem 7.12.

3.3 How Cardinality and Energy Interact

Cauchy-Schwarz inequality

We can quantify the relationship between the arithmetic energy and the cardinality of the sum or product set by using the Cauchy-Schwarz inequality.

Before we state this, we introduce another equivalent definition of the additive energy of a set $A \subseteq \mathbb{G}$, for $\mathbb{G}$ an abelian group.

Let $r_{A-B}(x) := |\{(a, b) \in A \times B \colon x = a - b\}|$ be the number of representations of x as an element of $A - B$. Then

$$\begin{aligned}\mathsf{E}^+(A, B) &= |\{(a, b, c, d) \in A \times B \times A \times B \colon a - b = c - d\}| \\ &= |\{(a, b, c, d, x) \in A \times B \times A \times B \times \mathbb{G} \colon a - b = c - d = x\}| \\ &= \sum_{x \in \mathbb{G}} r_{A-B}^2(x)\,.\end{aligned}$$

An analogous interpretation can be made for the multiplicative energy $\mathsf{E}^\times(A, B)$.

Lemma 3.11. *Let $A, B \subseteq \mathbb{G}$ be finite sets is an abelian group $\mathbb{G}$. Then*

$$|A|^2|B|^2 \leq \mathsf{E}^+(A, B)|A \pm B|\,. \tag{3.3}$$

Proof. We have

$$|A|^2|B|^2 = \left(\sum_{x\in\mathbb{G}} r_{A\pm B}(x)\right)^2 \leq \sum_{x\in A\pm B} 1^2 \sum_{x\in A\pm B} r^2_{A\pm B}(x) = |A\pm B|\ \mathsf{E}^+(A,B)\,.$$

The inequality arises from an application of the Cauchy-Schwarz inequality. □

For $A, B \subseteq \mathbb{F}$, the multiplicative version of Lemma 3.11 is the statement

$$|A|^2|B|^2 \leq \mathsf{E}^\times(A,B)|AB|\,. \tag{3.4}$$

We have a similar relation with the product set $A/B := \{a/b\colon a\in A, b\in B\setminus\{0\}\}$, except now we must be careful when $0\in B$. We leave this technicality to the reader.

We use Hölder's inquality (in place of the Cauchy-Schwarz inequality of Lemma 3.11) to obtain an analogue of (3.3) which connects the k-th energy with the difference set:

$$|A|^{2k} \leq \mathsf{E}^+_k(A)\ |A-A|^{k-1}\,.$$

Indeed, let l satisfy $\frac{1}{k}+\frac{1}{l}=1$. Then

$$|A|^2 = \sum_{x\in\mathbb{G}} r_{A-A}(x) \leq \left(\sum_{x\in A-A} 1^l\right)^{\frac{1}{l}} \left(\sum_x r^k_{A-A}(x)\right)^{\frac{1}{k}} = |A-A|^{\frac{k-1}{k}}\mathsf{E}^+_k(A)^{\frac{1}{k}}\,.$$

Addendum: The Balog-Szemerédi-Gowers Theorem

Although we do not use it in this thesis, the Balog-Szemerédi-Gowers theorem [3, 38, 95] is an important result in additive combinatorics. It is frequently used as a powerful tool in the literature, but also serves as an apparatus to develop the intangible concept of intuition in this area.

Theorem 3.12 (Balog-Szemerédi-Gowers Theorem)**.** *Let $A\subseteq\mathbb{G}$ for $\mathbb{G}$ an abelian group and let $K\geq 1$. If $\mathsf{E}^+(A)\geq |A|^3K^{-1}$ then there exists subsets $A'\subseteq A$ such that $|A'|\gg |A|K^{-1}$ and*

$$|A'-A'| \ll K^4|A'|\,.$$

This statement of the theorem is by Schoen [95], who has proved the strongest quantitative bounds in this direction; he also proves an asymmetric version.

The Balog-Szemerédi-Gowers theorem is an inverse theorem; if a set has a large energy, then a large subset has structure. As a quantitative tool, it is useful only when K is very small.

Solymosi's Approach to the Sum-Product Problem

Having now developed the necessary energetic terminology, we demonstrate the claimed elegance of Solymosi's approach to the sum-product problem. Solymosi [104] connected the multiplicative energy of a set with the cardinality of its sumset. Using the Cauchy-Schwarz inequality yields the following result.

Theorem 3.13. *Let $A \subseteq \mathbb{R}$ be a finite set. Then*

$$\mathsf{E}^{\times}(A) \lesssim |A+A|^2 . \tag{3.5}$$

Consequently

$$|A+A|^2|AA| \gtrsim |A|^4 \tag{3.6}$$

and

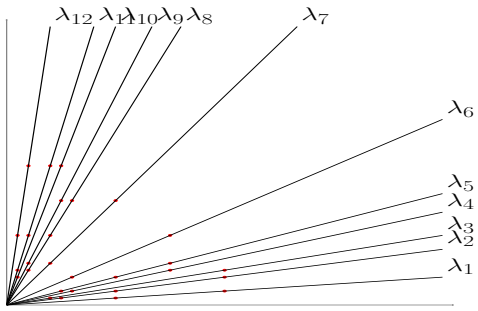
$$\max(|A+A|, |AA|) \gtrsim |A|^{4/3} . \tag{3.7}$$

Although we hide constants and logarithmic terms within the notation here, Solymosi calculates them explicitly.

Solymosi's proof proceeds as follows. First, without loss of generality, suppose $A \subseteq \mathbb{R}_{>0}$.[2] Consider the set of points $A \times A$ and the set of rays through the origin which intersect these points. Each line has slope in the set A/A. By a dyadic pigeonholing argument, we can assume that each ray has approximately (i.e. up to a factor of 2) the same number of points (say, t) in $A \times A$ on it.

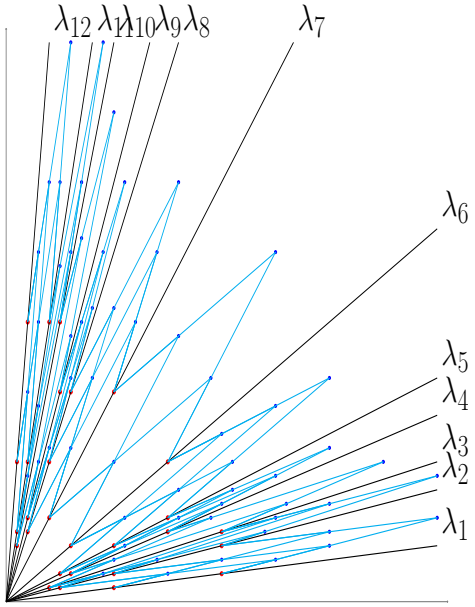
[2]This is justified by keeping at least half of the elements of A of the same sign; if these elements happen to be negative, set $B = -A$ and notice that $|B+B| = |A+A|$ and $|BB| = |AA|$.

Figure 3.3: Lines through the origin each contain (approximately) the same number of points of $A \times A$. We use the set $A = \{1, 2, 4, 5, 6, 10, 15, 20\}$. Then, with $t = 2$, we regularise so that each slope contains between t and $2t$ points



Suppose the slopes are ordered in a clockwise (or anti-clockwise – it does not matter) manner and consider two consecutive slopes corresponding to the lines $y = \lambda_i x$ and $y = \lambda_{i+1} x$. By thinking of the points in $A \times A$ lying on the lines with these slopes as vectors, observe that the vector sum of $(x, \lambda_i x) \in A \times A$ and $(y, \lambda_{i+1} y) \in A \times A$ lies in the set $(A + A) \times (A + A)$.

Figure 3.4: From every pair of consecutive slopes λ_i and λ_{i+1} we create elements of $(A + A) \times (A + A)$ by calculating the vector sum of a point on λ_i and a point on λ_{i+1}.



Moreover, these new vector sums, created only from pairs of vectors lying

on two consecutive slopes, do not coincide. Since every line of the form $y = \lambda_i x$ contains approximately (that is, up to a factor of 2) t points of $A \times A$, consecutive slopes generate t^2 new elements of $(A+A)\times(A+A)$. To be precise, consecutive slopes generate between t^2 and $4t^2$ elements of $(A + A) \times (A + A)$. The key points is that we create at least $t^2|A/A|/\log(|A|)$ many distinct vectors in $(A + A) \times (A + A)$, since there are at least $|A/A|/\log(|A|)$ slopes with supporting about t points of $A \times A$.

Finally, we relate this quantity to the multiplicative energy of A by the previously discussed interpretation of multiplicative energy. The multiplicative energy can be calculated as the sum, over the slopes in A/A, of the number of pairs of points of $A \times A$ lying on a slope. The number of new elements of $(A + A) \times (A + A)$ that we generate is less than this sum.

Hence we have that

$$|A + A|^2 \geq t^2|A/A|/\log(|A|) \sim \mathsf{E}^{\times}(A)\,.$$

Equations (3.6) follows directly from (3.5) via the Cauchy-Schwarz inequality (3.3) since $\mathsf{E}^{\times}(A) \geq |A|^4/|AA|$.

3.4 Other approaches to the sum-product phenomenon

The sum-product phenomenon is a statement about the impossibility of the coexistence of multiplicative and additive structure within a set. We present yet another direction to approach the notion guiding Conjecture 3.4.

Few sums or few products

We could already find ourselves in possession of a quantitative characterisation of a set – e.g. that the sum set is small. Then our task becomes to show that the product set is large. This 'few sums, many products approach' is particularly fruitful over $\mathbb{C}$, where Elekes and Ruzsa [26] used the Szemerédi-Trotter theorem to prove that

$$|A + A|^4|AA| \gg \frac{|A|^6}{\log|A|}\,. \tag{3.8}$$

Equation (3.8) is a consequence of counting the number of collinear triples in the point set $((A + A) \cup A) \times ((A + A) \cup A)$ – see Theorem 2.12. Elekes and

Ruzsa used the observation that if $(a,b,c,d) \in A^4$ is such that $ab = cd$, then for any $e, f \in A$, the points (e,f), $(e+a, f+c)$, $(d+e, b+f)$ form a collinear triple.

The converse problem of 'few products, many sums' is harder. Over the rationals, the strongest result is due to Bourgain and Chang [15] who show that if $|AA| = M|A|$, then $|A+A| \gg M^{-C(\epsilon)}|A|^{2-\epsilon}$ where $C(\epsilon) \to \infty$ as $\epsilon \to 0$. Over the complex numbers, much less is known. The state-of-the-art in this direction is a result of Olmezov, Semchankau and Shkredov [77], who recently proved that

$$|A+A|^{10}|AA|^{14} \gtrsim |A|^{30}\,.$$

They proved stronger results when the sum set is replaced by the difference set $A - A$.

On the size of $|AA + A|$ or $|A(A+A)|$

The sum-product problem is to show that at least one of two sets is large; a weaker variant is to show that a single set combining multiplication and addition must grow.

Popular candidates for this variant are the sets

$$A + AA := \{a + bc : a, b, c \in A\} \text{ and } A(A+A) := \{a(b+c) : a, b, c \in A\}\,.$$

If we take the examples of A being an arithmetic or geometric progression, it is clear that we cannot hope for $|A(A+A)|$ or $|A+AA|$ to be larger than $|A|^2$ in cardinality (in fact, because of the multiplication table problem, presented in Example 3.2, $|A+AA|, |A(A+A)|$ can be $o(|A|^2)$).

Proving the growth of the objects $A(A+A)$ or $AA+A$ would be strong evidence towards the sum-product conjecture.

Using just the Szemerédi-Trotter theorem, it is relatively simple to show that $|A(A+A)|, |AA+A| \gg |A|^{3/2}$. Passing beyond this threshold is an area of active research with the best results currently as follows: Murphy, Roche-Newton and Shkredov [71] show that $|A(A+A)| \gg |A|^{3/2+1/186}$; Roche-Newton and Warren [87] show that $|AA+A| \gg |A|^{3/2+1/194}$.

4

Incidences between points and lines in arbitrary fields

4.1 Introduction

Let $\mathcal{P}$ be a finite set of points in $\mathbb{F}^2$, where $\mathbb{F}$ is a field, and let $\mathcal{L}$ be a finite set of lines in $\mathbb{F}^2$.

In this chapter, we will study $\mathcal{I}(\mathcal{P}, \mathcal{L})$, the number of incidences between $\mathcal{P}$ and $\mathcal{L}$. We first review what is known over $\mathbb{R}$, and then we consider point and line sets over arbitrary fields. The main new content of this chapter will be two types of new incidence bounds:

1. The first incidence bound is between a set of points $\mathcal{P} = A \times B$ and lines $\mathcal{L}$. This bound is optimal for certain families points and lines.

2. In the second incidence bound, we do not require that the point set has a Cartesian product structure, and we prove an incidence bound by bootstrapping the first incidence bound.

The content of this chapter is joint work with Frank de Zeeuw and appears in the publication [108]. This paper originally used a weaker form of the forthcoming Theorem 4.5, an incidence bound between points $A \times B$ and lines $\mathcal{L}$, to develop an incidence bound between unstructured sets of points and lines. De Zeeuw realised that Rudnev's point plane incidence theorem (Theorem 2.9) could be more efficiently exploited to yield the optimal Cartesian-Product incidence bound that is presented here as Theorem 4.4. The proofs appear in [108] – any significant text overlap is restricted to proofs that I initially wrote up.

4.2 Incidences over $\mathbb{R}$

We recall the Szemerédi-Trotter bound from Chapter 2; for finite sets of points $\mathcal{P}$ and lines $\mathcal{L}$ in $\mathbb{R}^2$ the number of incidences between $\mathcal{P}$ and $\mathcal{L}$ is bounded by

$$\mathcal{I}(\mathcal{P}, \mathcal{L}) \leq c_{ST} \left((|\mathcal{P}||\mathcal{L}|)^{2/3} + |\mathcal{P}| + |\mathcal{L}| \right) .$$

A particularly insightful proof of the Szemerédi-Trotter theorem is via polynomial partitioning, a method developed by Guth and Katz (see e.g. [43] and references therein). In this later proof, often attributed to Tao [112], a relatively low-degree polynomial exists that divides the plane into *cells* – open sets each containing (approximately) the same number of points. Lemma 2.1 then counts the number of incidences between the points in the cell and lines passing through each cell. Applying the trivial bound in this manner produces a stronger bound since we are using it in a more optimal way, having reduced to a more uniform situation. It then remains to count the contribution of incidences coming from points on the cell wall. The key observation here is that a given line cannot pass through too many cell walls, an idea which uses the topology of the real plane.

The original proof by Szemerédi and Trotter [110] was far more computationally involved, and involves successive refinements of the set of points into increasingly regular subsets. This idea is a precursor to the *cell partitioning* method introduced by Clarkson et al. [23]. The cell partitioning method is a way of dividing points into cells without the modern machinery of the polynomial method.

Another 'textbook proof' of Theorem 2.3 is that of Székely [109] which uses the Crossing Lemma. This relates the configuration of points and lines to a planar graph; this technique has yielded the most success in attempts to calculate the exact constant c_{ST}. (The original proof of Theorem 2.3 showed that one could take $c_{ST} = 10^{60}$.) Currently the world record is held by Ackerman [1, Corollary 3.3]

$$\mathcal{I}(\mathcal{P}, \mathcal{L}) \leq 2.44|\mathcal{P}|^{2/3}|\mathcal{L}|^{2/3} + |\mathcal{P}| + |\mathcal{L}| .$$

Both the cell partitioning method and the crossing lemma method used to prove the Szemerédi-Trotter bound rely on special properties of $\mathbb{R}$; the cell decomposition method uses the topology of the reals whilst the crossing lemma relies on the topological Euler characteristic of the plane.

4.3 Incidences in Finite Fields

We recall the obstruction presented in Chapter 2 to an incidence bound in an arbitrary field: let $\mathcal{P}$ be the set of all points in $\mathbb{F}_p^2$ and let $\mathcal{L}$ be the set of all lines in $\mathbb{F}_p^2$, then $\mathcal{I}(\mathcal{P}, \mathcal{L}) \gg |\mathcal{P}|\sqrt{|\mathcal{L}|}$, which coincides asymptotically with the trivial bound attained by Lemma 2.1. However, if we remove this example from the sets of points and lines that we consider, then we are able to find a non-trivial incidence bound in finite fields.

The first work in this direction was by Bourgain, Katz and Tao [17] in 2003: they established a non-trivial incidence bound in finite fields of prime order.

Theorem 4.1 (Bourgain, Katz, Tao [17])**.** *Let $\mathbb{F} = \mathbb{F}_p$ for a prime p. Let $\mathcal{P}$ be a finite set of points and $\mathcal{L}$ a finite set of lines in $\mathbb{F}^2$ with $|\mathcal{P}|, |\mathcal{L}| \leq N = p^\alpha$ for some $0 < \alpha < 2$. Then there exists an absolute constant $C \geq 0$ such that:*

$$\mathcal{I}(\mathcal{P}, \mathcal{L}) \leq CN^{3/2-\epsilon}$$

for some $\epsilon = \epsilon(\alpha) > 0$.

They remark also that a non-trivial incidence bound exists in more general fields as long as the point set $\mathcal{P}$ does not have large intersection with (a projective transformation of) $\mathbb{G} \times \mathbb{G}$, where $\mathbb{G}$ is a subfield of $\mathbb{F}$, and the line set $\mathcal{L}$ does not have large intersection with the associated collection of lines. That is, when we dualise the line set $\mathcal{L}$ to obtain a collection of points, these points do not have large intersection with $\mathbb{G} \times \mathbb{G}$.

They achieved this by first proving a non-trivial sum-product bound using both delicate and difficult combinations of elementary tools of additive combinatorics, and more powerful combinatorial tools such as the Balog-Szemerédi-Gowers Theorem discussed in Section 3.3. From a non-trivial sum-product bound, they extrapolated a non-trivial qualitative incidence bound.

An explicit value $\epsilon = 1/10678$ for $\alpha = 1$ was found by Helfgott and Rudnev [46], and further improvements to ϵ appeared in the work of Jones [50, 51], with the best (prior to this work) bound summarised below.

Theorem 4.2. (Jones [51]) *Let $\mathcal{P}$ be a finite set of points in $\mathbb{F}_p^2$ and $\mathcal{L}$ a finite set of lines in $\mathbb{F}^2$, with $|\mathcal{P}|, |\mathcal{L}| \leq N < p$. Then we have*

$$\mathcal{I}(\mathcal{P}, \mathcal{L}) \ll N^{3/2-\epsilon},$$

with $\epsilon = 1/662$.

Jones originally stated his result only over $\mathbb{F}_p$, as his proof relied on a sum-product-type energy inequality in $\mathbb{F}_p$. It is a short calculation to show that his bound improves to $\epsilon \geq 1/326$ using recent bounds. This follows by replacing [51, Lemma 11] in Jones's proof by a suitable multiplicative analogue of [86, Theorem 6]. Moreover, the application of [86] swiftly extends his result to any field, albeit with the restriction $N < p$ in terms of the *characteristic* of the field, instead of the cardinality of $\mathbb{F}$.

Prior to the new work contained in this thesis, a stronger incidence bound was known if the point set is a Cartesian product. We remark that by duality, this bound is also applicable if the line set has a Cartesian product structure. That is, the line set $\mathcal{L}$ contains lines of the form

$$\ell_{ab} := \{(x, y) \in \mathbb{F}^2 \colon y = ax + b\}$$

and $\mathcal{L} = \{\ell_{ab} \colon (a, b) \in A \times B\}$.

Theorem 4.3 (Aksoy Yazici et al. [118])**.** *Let $\mathbb{F}$ be a field with characteristic $p > 0$. Let $A \subset \mathbb{F}$ be such that $|A| \ll p^{2/3}$ and let $\mathcal{P} = A \times A$ be a finite point set with $|\mathcal{P}| = |A|^2$. Let $\mathcal{L} \subseteq \mathbb{F}^2$ be a finite set of lines. Then*

$$\mathcal{I}(\mathcal{P}, \mathcal{L}) \ll |\mathcal{P}|^{3/4}|\mathcal{L}|^{2/3} + |\mathcal{L}|\,.$$

The proof of this theorem follows from a count on the number of collinear triples, which in turn relies on an incidence bound of Rudnev [89] between points and planes (Theorem 2.9.

The original manuscript of the note [108] used Theorem 4.3 to obtain an incidence bound between arbitrary (finite) sets of points and lines in any field (under suitable constraints on $|\mathcal{P}|, |\mathcal{L}|$ in terms of the characteristic of the field). This bound is improved in the forthcoming Theorem 4.5. To pass from an incidence bound between lines and points in a Cartesian product formation to an incidence bound between arbitrary sets of lines and points, we use a bootstrapping technique. Roughly speaking, this involves finding many 'large grids' of points that, under a projective transformation, become Cartesian products.

4.4 Main Results

Theorem 4.4. *[Szemerédi-Trotter in Arbitrary Fields] Let $\mathcal{P}$ be a finite set of points in $\mathbb{F}^2$ and $\mathcal{L}$ a finite set of lines in $\mathbb{F}^2$, with $|\mathcal{P}|^{7/8} < |\mathcal{L}| < |\mathcal{P}|^{8/7}$. If $\mathbb{F}$*

has positive characteristic p, assume in addition that $|\mathcal{P}|^{-2}|\mathcal{L}|^{13} \ll p^{15}$. *Then*

$$\mathcal{I}(\mathcal{P}, \mathcal{L}) \ll |\mathcal{P}|^{11/15}|\mathcal{L}|^{11/15}. \tag{4.1}$$

Theorem 4.4 is stronger than the trivial bound Lemma 2.1 when $|\mathcal{P}|^{7/8} < |\mathcal{L}| < |\mathcal{P}|^{8/7}$. We could choose to present the theorem without this condition, in which case equation (4.1) would be replaced by the bound

$$\mathcal{I}(\mathcal{P}, \mathcal{L}) \ll |\mathcal{P}|^{11/15}|\mathcal{L}|^{11/15} + |\mathcal{P}| + |\mathcal{L}|.$$

When the point set is a Cartesian product, we have the following strengthening:

Theorem 4.5. *[Cartesian Szemerédi-Trotter] Let* $\mathcal{L}$ *be a finite set of lines in* $\mathbb{F}^2$. *Let* $A, B \subset \mathbb{F}$ *be sets with* $|A| \leq |B|$ *and* $|A||B|^2 \leq |\mathcal{L}|^3$. *If* $\mathbb{F}$ *has positive characteristic p, assume that* $|A||\mathcal{L}| \ll p^2$. *Then*

$$\mathcal{I}(A \times B, \mathcal{L}) \ll |A|^{3/4}|B|^{1/2}|\mathcal{L}|^{3/4} + |\mathcal{L}|.$$

As before, changing the condition $|A||\mathcal{L}| \ll p^2$ to $|A||\mathcal{L}| \leq p^2$ changes only the constant in the subsequent incidence estimate.

In the special case of $|A| = |B| = \sqrt{|\mathcal{P}|}$, this gives the bound

$$\mathcal{I}(A \times A, \mathcal{L}) \ll |\mathcal{P}|^{5/8}|\mathcal{L}|^{3/4} + |\mathcal{L}|.$$

We note that Theorem 4.5 has been used to provide an optimal bound on the number of collinear quadruples in a set. Petridis [79, 70], showed that the number of collinear quadruples in $A \times A$, denoted $Q(A)$ satisfies

$$Q(A) = O\left(\frac{|A|^8}{p^2} + \log(|A|)|A|^5\right). \tag{4.2}$$

This count is optimal (up to logarithmic factors and constants). We provide the proof of an asymmetric version of (4.2) as Theorem 5.7.

4.5 Discussion

Optimality

Theorem 4.5 is quantitatively weaker than the Szemerédi-Trotter bound of Theorem 2.3 . For instance, if we consider Cartesian products $\mathcal{P} = A \times A$ with

$|\mathcal{L}| = |\mathcal{P}| = N$, then Theorem 4.5 gives the bound $\mathcal{I}(\mathcal{P}, \mathcal{L}) \ll N^{11/8}$, whereas Theorem 2.3 gives $\mathcal{I}(\mathcal{P}, \mathcal{L}) \ll N^{4/3}$ (over $\mathbb{C}$). However, Theorem 4.5 is optimal for certain sets of points and lines, as the following example demonstrates. The example closely follows a well-known construction of Elekes [25] over the reals.

Example 4.6. *Let $0 < s \leq t$ be integers such that $st = O(p)$.*

Let $I = \{1, \ldots, s\}$ and let $J = \{1, \ldots, 2st\}$ be subsets of $\mathbb{F}$. Let $\mathcal{P} = I \times J \subset \mathbb{F}^2$ be a set of points and let $\mathcal{L}$ be the set of lines of the form $y = mx + c$ where $m \in \{1, \ldots, t\}$ and $c \in \{1, \ldots, st\}$.

We have $2s^2t$ points and st^2 lines in $\mathbb{F}^2$. By construction, every line in $\mathcal{L}$ contains s points of $\mathcal{P}$ and so $\mathcal{I}(\mathcal{P}, \mathcal{L}) = s^2t^2$. Theorem 4.5 yields the matching bound

$$\mathcal{I}(\mathcal{P}, \mathcal{L}) \ll s^{3/4}(2st)^{1/2}(st^2)^{3/4} \ll s^2t^2 \,.$$

However, this example, although optimal, is somewhat contrived, and it seems likely that Theorem 4.5 is not optimal for all choices of $|A|$, $|B|$ and $|L|$.

We do not believe that the exponent in Theorem 4.4 of $11/15$ is optimal. This belief stems from the analogous situation in the reals: all known examples of optimality of the Szemerédi-Trotter theorem depend on variations of a lattice structure.

Bounding rich points and lines

Recall that a point p is said to be k-rich with respect to a line configuration $\mathcal{L}$ if $|\{\ell \in \mathcal{L} \colon p \in \ell\}| \geq k$; similarly, a line ℓ is said to be k-rich with respect to a set of points $\mathcal{P}$ if $|\{p \in \mathcal{P} \colon p \in \ell\}| \geq k$.

Theorem 4.7 (Rich Lines in $\mathcal{P}$)**.** *Let $\mathcal{P} \subseteq \mathbb{F}^2$ be a finite set of points and let $\sqrt{|\mathcal{P}|} \ll k \leq |\mathcal{P}|$ be an integer. Let $\mathcal{L}_k$ be the set of k-rich lines with respect to $\mathcal{P}$.*

In positive characteristic p, assume that $|\mathcal{P}|^{11}k^{-13} \leq p^{15}$.

Then

$$|\mathcal{L}_k| \ll \frac{|\mathcal{P}|^{11/4}}{k^4} + \frac{|\mathcal{P}|}{k} \,.$$

In the statement of the theorem we assume that $k \gg \sqrt{|\mathcal{P}|}$; otherwise we use the trivial incidence bound Lemma 2.1.

Proof. By the trivial incidence bound Lemma 2.1, we can crudely bound $|\mathcal{L}_k|$ as $|\mathcal{L}_k| \ll \min(|\mathcal{P}|k^{-1}, |\mathcal{P}|^2k^{-2})$. In the subsequent, both these bounds are utilised.

Note that since $k \gg 1$, we must have that $|\mathcal{L}_k| \ll |\mathcal{P}|^{8/7}$. This follows from the trivial bound.

If $|\mathcal{L}_k| \ll |\mathcal{P}|^{7/8}$, we apply the trivial bound to obtain $|\mathcal{L}_k| \ll |\mathcal{P}|k^{-1}$. This bound is better than the inequality $|\mathcal{L}_k| \ll |\mathcal{P}|^{7/8}$ when $k \gg |\mathcal{P}|^{1/8}$.

In order to apply Theorem 4.4, we require that $|\mathcal{P}|^{-2}|\mathcal{L}_k|^{13} \ll p^{15}$. This follows from the trivial bound and the hypotheses of the theorem. Applying Theorem 4.4 completes the proof. □

We could of course prove Theorem 4.7 using the trivial incidence bound Lemma 2.1 in place of the stronger Theorem 4.4; this crude bound yields $|\mathcal{L}_k| \ll |\mathcal{P}|^2k^{-2}$. A comparison of these two bounds tells us that Theorem 4.7 is better than the trivial bound whenever $k > |\mathcal{P}|^{3/8}$ and $k < |\mathcal{P}|^{7/12}$.

In the case when one of the set of points or lines has a Cartesian product structure, we have a stronger bound from Theorem 4.5.

Theorem 4.8 (Rich lines in $A \times B$)**.** *Let $A \times B \subseteq \mathbb{F}^2$ be a finite set of points with $|A| \leq |B|$ and let $|A|^{1/2} \leq k \leq |A|^{1/2}|B|^{1/2} + |A|^{2/3}|B|^{1/3}$ be an integer. Let $\mathcal{L}_k$ be the set of k-rich lines with respect to $A \times B$.*

In positive characteristic p, assume that $\frac{|A|^3|B|^2}{k^2} \leq p^2$.

Then

$$|\mathcal{L}_k| \ll \frac{|A|^3|B|^2}{k^4}\,.$$

Proof. Although the proof is identical to that of Theorem 4.7, the constraints on k are more involved.

Firstly, note that by the trivial bound we have

$$|\mathcal{L}_k| \leq \min\left(\frac{|A||B|}{k}, \frac{|A|^2|B|^2}{k^2}\right)\,.$$

The second minimand and the conditions in the statement of the theorem ensures that $|\mathcal{L}_k||A| \ll p^2$, which is the condition we must satisfy to apply Theorem 4.5.

Applying the Cartesian incidence bound of Theorem 4.5, we obtain that either $|A||B|^2 > |\mathcal{L}_k|^3$ or

$$k|\mathcal{L}_k| \ll \mathcal{I}(A \times B, \mathcal{L}_k) \ll |A|^{3/4}|B|^{1/2}|\mathcal{L}_k|^{3/4} + |\mathcal{L}_k|\,.$$

Since $k \gg 1$, the second term is subsumed by the first term and the implicit constants hidden in the notation.

Hence $|\mathcal{L}_k| \ll \frac{|A|^3|B|^2}{k^4}$. Note that this is a smaller bound than $|\mathcal{L}_k|^3 < |A||B|^2$.

This bound is better than the trivial bound from Lemma 2.1 for the range of k in the statement of the theorem. □

Both Theorem 4.7 and Theorem 4.8 pertain to rich lines with respect to a point set. By duality, we can exchange the roles of points and lines to obtain a bound on the number of k-rich points relative to a line set.

When to use Theorems 4.4 and 4.5?

Theorem 4.5 is always better than Theorem 4.4, and should be used in preference whenever some sort of Cartesian product structure exists.

However, although Cartesian products naturally arise in applications of the incidence bound, there are many occasions to use Theorem 4.4. When $|\mathcal{P}| = |\mathcal{L}| = N$, Theorem 4.4 improves the ϵ in Theorem 4.2 from 1/662 to 1/30, it extends the condition in positive characteristic to $N \ll p^{15/11}$, and it has the further advantage of being sensitive to the relative sizes of the point set and line set. To compare it with the bound of Vinh [116], assume $\mathbb{F} = \mathbb{F}_p$ and $m = n = N$; then Theorem 4.4 is better for $N \ll p^{15/14}$.

However, outside the range $|\mathcal{P}|^{7/8} < |\mathcal{L}| < |\mathcal{P}|^{8/7}$ the bounds from Lemma 2.1 are better. Theorem 4.4 is always better than Jones' Theorem 4.2.

We summarise the situation in Table 4.5.

Range of $\lvert\mathcal{L}\rvert$	Best bound
$\lvert\mathcal{L}\rvert < \lvert\mathcal{P}\rvert^{1/2}$	$\ll \lvert\mathcal{P}\rvert$
$\lvert\mathcal{P}\rvert^{1/2} < \lvert\mathcal{L}\rvert < \lvert\mathcal{P}\rvert^{7/8}$	$\ll \lvert\mathcal{P}\rvert^{1/2}\lvert\mathcal{L}\rvert$
$\lvert\mathcal{P}\rvert^{7/8} < \lvert\mathcal{L}\rvert < \lvert\mathcal{P}\rvert^{8/7}$	$\ll \lvert\mathcal{P}\rvert^{11/15}\lvert\mathcal{L}\rvert^{11/15}$
$\lvert\mathcal{P}\rvert^{8/7} < \lvert\mathcal{L}\rvert < \lvert\mathcal{P}\rvert^{2}$	$\ll \lvert\mathcal{P}\rvert\lvert\mathcal{L}\rvert^{1/2}$
$\lvert\mathcal{P}\rvert^{2} < \lvert\mathcal{L}\rvert$	$\ll \lvert\mathcal{L}\rvert$

Table 4.1: Overview of best known upper bounds on $\mathcal{I}(\mathcal{P}, \mathcal{L})$

We note that if we have point and line sets in $\mathbb{F}_q$, with q a prime power, with cardinalities (quantifiably) *large* in terms of the characteristic, then, in

certain situations, recent results of Mohammadi [66, Theorems 1 and 2] provide the best known incidence bound for the situation.

The rest of this chapter is devoted to proving Theorems 4.4 and 4.5.

4.6 Proof of Theorem 4.5: Cartesian incidence bound

Rudnev's bound, Theorem 2.9, together with the Cauchy-Schwarz inequality and some knowledge of projective geometry are the only tools required for Theorem 4.5. The necessary projective geometry background is contained in Section 2.4.

We begin the proof of Theorem 4.5 by removing all vertical lines from $\mathcal{L}$; these contribute at most $|A||B|$ incidences. We use the assumption $|A||B|^2 \leq |\mathcal{L}|^3$ to show that these incidences contribute less than $|A|^{3/4}|B|^{1/2}|\mathcal{L}|^{3/4}$ to the count. We remove vertical lines in order to be able to dualise the lines.

Secondly, we can assume that $|B|^2 \leq |A||\mathcal{L}|$. Indeed, having removed all vertical lines, we have the bound $\mathcal{I}(A \times B, \mathcal{L}) \leq |A||\mathcal{L}|$. If $|B|^2 > |A||\mathcal{L}|$, then we have that $\mathcal{I}(A \times B, \mathcal{L}) \leq |A||\mathcal{L}| \leq |A|^{3/4}|B|^{1/2}|\mathcal{L}|^{3/4}$ and so we are done. This assumption is made in order to legitimise our third and final assumption.

Thirdly, we claim that at most $|A|^{1/2}|\mathcal{L}|^{1/2}$ lines of $\mathcal{L}$ are concurrent or parallel. Indeed, we iteratively remove any pencil of more than $|A|^{1/2}|\mathcal{L}|^{1/2}$ concurrent or parallel lines. (In projective space, parallel lines intersect at infinity, so parallel lines are concurrent in projective geometry.) Let n_i be the number of lines in the i-th pencil that we remove (not counting those that were removed earlier). Then the i-th pencil is involved in at most $|A||B| + n_i$ incidences. We remove all pencils in at most $|\mathcal{L}|/(|A|^{1/2}|\mathcal{L}|^{1/2}) = |A|^{-1/2}|\mathcal{L}|^{1/2}$ iterations, discounting at most

$$|A|^{-1/2}|\mathcal{L}|^{1/2}|A||B| + \sum n_i \leq |A|^{1/2}|B||\mathcal{L}|^{1/2} + |\mathcal{L}|$$

incidences.

We note here that, by the assumption $|B|^2 \leq |A||\mathcal{L}|$, the quantity above is at most $|A|^{3/4}|B|^{1/2}|\mathcal{L}|^{3/4} + |\mathcal{L}|$. Hence we have a bound on the maximum number of concurrent or parallel lines.

Having removed lines in this manner, we will abuse notation, and in the subsequent, denote this reduced set of lines as $\mathcal{L}$. We consider the affine dual

of $\mathcal{L}$, which is now well-defined as we have removed all vertical lines from $\mathcal{L}$. Let

$$\mathcal{L}^* := \{(c,d) \in \mathbb{F}^2 : y = cx + d \in \mathcal{L}\}$$

Then we have

$$\begin{aligned}\mathcal{I}(\mathcal{P},\mathcal{L}) &= |\{(x,y,s,t) \in A \times B \times \mathcal{L}^* : xs + t = y\}| \\ &= \sum_{y \in B} \sum_{(x,s,t) \in A \times \mathcal{L}^*} \mathbb{1}_{xs+t=y} \\ &\leq |B|^{1/2} X^{1/2},\end{aligned}$$

where

$$X := |\{(x,s,t,x',s',t') \in A \times \mathcal{L}^* \times A \times \mathcal{L}^* : xs + t = x's' + t'\}|.$$

The above inequality is an application of Cauchy-Schwarz in y. We bound the quantity X using Rudnev's point-plane incidence bound (Theorem 2.9).

Define a point set and a plane set by

$$\mathcal{R} := \{(x,s',t') \in A \times \mathcal{L}^*\}, \qquad \mathcal{S} := \{xs + t = x's' + t' : (x',s,t) \in A \times \mathcal{L}^*\}.$$

We have $|\mathcal{R}| = |\mathcal{S}| = |A||\mathcal{L}|$ and $X = \mathcal{I}(\mathcal{R},\mathcal{S})$.

To apply Theorem 2.9 we need to check its conditions. The condition that the number of points is $O(p^2)$ follows from the assumption that $|A||\mathcal{L}| \ll p^2$. The condition that there are at most as many points as planes clearly holds, since $|\mathcal{R}| = |\mathcal{S}|$. Because of the product structure of $\mathcal{R} = A \times \mathcal{L}^*$, the maximum number of collinear points in $\mathcal{R}$ is bounded by the maximum of $|A|$ and the maximum number of collinear points in $\mathcal{L}^*$. The former is bounded by $|A|^{1/2}|\mathcal{L}|^{1/2}$, using the fact that $|A| \leq |\mathcal{L}|$, which follows since $|A| \leq |B|$ and $|A||B|^2 \leq |\mathcal{L}|^3$. The maximum number of collinear points in $\mathcal{L}^*$ equals (by duality) the maximum number of concurrent lines in $\mathcal{L}$. This is bounded by $|A|^{1/2}|\mathcal{L}|^{1/2}$ by our initial refinement of $\mathcal{L}$.

We can apply Theorem 2.9 with $k = |A|^{1/2}|\mathcal{L}|^{1/2}$ to obtain

$$\mathcal{I}(\mathcal{R},\mathcal{S}) \ll |\mathcal{R}|^{1/2}|\mathcal{S}| + k|\mathcal{S}| \ll |A|^{3/2}|\mathcal{L}|^{3/2}.$$

Thus

$$\begin{aligned}\mathcal{I}(A \times B,\mathcal{L}) &\ll (|A|^{3/4}|B|^{1/2}|\mathcal{L}|^{3/4} + |\mathcal{L}|) + |B|^{1/2}X^{1/2} \\ &\ll |A|^{3/4}|B|^{1/2}|\mathcal{L}|^{3/4} + |\mathcal{L}|,\end{aligned}$$

as required.

4.7 Proof of Theorem 4.4

In order to apply Theorem 4.5 to an unstructured point set, we require a means to find large grids in a point set with many incidences. This approach was first taken in the original incidence bound over $\mathbb{F}$ in [17], where the authors showed that if a point set has many incidences, then a large subset of the points can be captured inside the intersection of two relatively small pencils. Then they used the fact that the set of intersection points of two pencils is projectively equivalent to a Cartesian product. This approach was quantitatively refined by Jones in [51], who showed that, after carefully 'regularising' the points, $\mathcal{P}$ can be efficiently partitioned into a number of subsets, each of which is covered by two relatively small pencils.

Our approach is also based on the fact that if a set is 'regular' in the sense that each point lies on a similar number of lines, then there are two pencils whose intersection covers many points of $\mathcal{P}$. This fact is captured in Lemma 4.9 below. This lemma is a quantitative version of Proposition 4 of [51]. We avoid asymptotic notation in this section, because in the next section we will apply Lemma 4.9 inside an induction, where we have to be careful with the dependence of the constants. We denote by $\overline{pq}$ the line in $\mathbb{F}^2$ containing the points $p, q \in \mathbb{F}^2$.

This proof of Theorem 4.4 (and of Lemma 4.9) also appears in the published version of the result [108].

Lemma 4.9. *The following holds for any constants $c_2 > c_1 > 0$.*

Let $\mathcal{P}$ be a set of m points and $\mathcal{L}$ a set of n lines, such that between c_1K and c_2K lines of $\mathcal{L}$ pass through each point in $\mathcal{P}$. Assume $K \geq 4n/(c_1m)$, $K \geq 8/c_1$ and $K^3 \geq 2^6n^2/(c_1^3m)$.

Then there are distinct points $p_1, q_1 \in \mathcal{P}$ and a set $G \subseteq \mathcal{P}\backslash\overline{p_1q_1}$ of cardinality $|G| \geq c_1^4K^4m/(2^9n^2)$, such that G is covered by at most c_2K lines from $\mathcal{L}$ through p_1, and by at most c_2K lines from $\mathcal{L}$ through q_1.

Proof. Let

$$\mathcal{L}_1 := \{\ell \in \mathcal{L} : |\ell \cap \mathcal{P}| \geq \mathcal{I}(\mathcal{P}, \mathcal{L})/(2n)\}.$$

Then we have $\mathcal{I}(\mathcal{P}, \mathcal{L}_1) \geq \mathcal{I}(\mathcal{P}, \mathcal{L})/2$, since the set of lines not contained in $\mathcal{L}_1$ contribute fewer than $n \cdot \mathcal{I}(\mathcal{P}, \mathcal{L})/(2n) = \mathcal{I}(\mathcal{P}, \mathcal{L})/2$ incidences to $\mathcal{I}(\mathcal{P}, \mathcal{L})$. Let $p_1 \in \mathcal{P}$ be a point incident to at least $\mathcal{I}(\mathcal{P}, \mathcal{L}_1)/(2m)$ lines in $\mathcal{L}_1$. Such a point

exists since the set of points that are incident to fewer than $\mathcal{I}(\mathcal{P}, \mathcal{L}_1)/2m$ lines contribute fewer than $m \cdot \mathcal{I}(\mathcal{P}, \mathcal{L}_1)/(2m) = \mathcal{I}(\mathcal{P}, \mathcal{L}_1)/2$ incidences to $\mathcal{I}(\mathcal{P}, \mathcal{L}_1)$.

Note that the assumptions of the lemma imply $\mathcal{I}(\mathcal{P}, \mathcal{L}) \geq c_1 K m$, so we have $\mathcal{I}(\mathcal{P}, \mathcal{L})/(2n) \geq c_1 K m/(2n)$ and

$$\mathcal{I}(\mathcal{P}, \mathcal{L}_1)/(2m) \geq (\mathcal{I}(\mathcal{P}, \mathcal{L})/2)/(2m) \geq c_1 K/4\,.$$

Thus the point p_1 is incident to at least $c_1 K/4$ lines from $\mathcal{L}_1$, and each line in $\mathcal{L}_1$ is incident to at least $(c_1 K m/(2n)) - 1$ points in $\mathcal{P}\backslash\{p_1\}$. It follows that

$$\mathcal{Q} := \{q \in \mathcal{P}\backslash\{p_1\} : \overline{p_1 q} \in \mathcal{L}\}$$

satisfies

$$|\mathcal{Q}| \geq \frac{c_1 K}{4}\left(\frac{c_1 K m}{2n} - 1\right) \geq \frac{c_1^2 K^2 m}{2^4 n}, \tag{4.3}$$

where in the last inequality we used the assumption $K \geq 4n/(c_1 m)$.

The points in $\mathcal{Q}$ still have the property that between $c_1 K$ and $c_2 K$ lines of $\mathcal{L}$ pass through them, so we can repeat the argument above, with $\mathcal{Q}$ in the role of $\mathcal{P}$, and the same line set $\mathcal{L}$. We let

$$\mathcal{L}_2 := \{\ell \in \mathcal{L} : |\ell \cap \mathcal{Q}| \geq \mathcal{I}(\mathcal{Q}, \mathcal{L})/(2n)\}\,.$$

As above, we have $\mathcal{I}(\mathcal{Q}, \mathcal{L}_2) \geq \mathcal{I}(\mathcal{Q}, \mathcal{L})/2$, and there is a point $q_1 \in \mathcal{Q}$ that is incident to at least $\mathcal{I}(\mathcal{Q}, \mathcal{L}_2)/(2|\mathcal{Q}|) \geq c_1 K/4$ lines in $\mathcal{L}_2$. Thus q_1 is incident to at least $(c_1 K/4) - 1$ lines from $\mathcal{L}_2$ other than the line $\overline{p_1 q_1}$, and each line in $\mathcal{L}_2$ is incident to at least $\mathcal{I}(\mathcal{Q}, \mathcal{L})/(2n) \geq c_1 K|\mathcal{Q}|/(2n)$ points in $\mathcal{P}$. Thus the set

$$\mathcal{R} := \{q \in \mathcal{Q}\backslash\overline{p_1 q_1} : \overline{q_1 q} \in \mathcal{L}_2\}$$

satisfies

$$|\mathcal{R}| \geq \left(\frac{c_1 K}{4} - 1\right)\left(\frac{c_1 K|\mathcal{Q}|}{2n} - 1\right) \geq \frac{c_1 K}{8} \cdot \frac{c_1 K|\mathcal{Q}|}{4n} = \frac{c_1^2 K^2 |\mathcal{Q}|}{2^5 n} \geq \frac{c_1^4 K^4 m}{2^9 n^2},$$

where in the second inequality we used $K \geq 8/c_1$ in the first factor, and both (4.3) and $K^3 \geq 2^6 n^2/(c_1^3 m)$ in the second factor, while in the last inequality we used (4.3).

As p_1 is incident to at most $c_2 K$ lines, $\mathcal{Q}$ is covered by at most $c_2 K$ lines from $\mathcal{L}$ that pass through p_1, and therefore so is $\mathcal{R} \subset \mathcal{Q}$. Similarly, $\mathcal{R}$ is covered by at most $c_2 K$ lines from $\mathcal{L}$ that pass through q_1. Therefore, we can choose the point set G as a subset of $\mathcal{R}$ with $|G| \geq c_1^4 K^4 m/(2^9 n^2)$. Note that we can simply take $G = \mathcal{R}$. This concludes the proof. □

Proof of Theorem 4.4

We will prove that there exists a constant C such that, for all $\mathcal{P}$ and $\mathcal{L}$ with $|\mathcal{L}|^{7/8} < |\mathcal{P}| < |\mathcal{L}|^{8/7}$, we have

$$I(\mathcal{P}, \mathcal{L}) < C|\mathcal{P}|^{11/15}|\mathcal{L}|^{11/15}.$$

Let $|\mathcal{P}| = m$ and $|\mathcal{L}| = n$. Our proof proceeds by induction, keeping n fixed and varying m. The inductive hypothesis is that for any point set $\mathcal{P}'$ satisfying $|\mathcal{P}'| = m'$, where $n^{7/8} < m' < m$, we have $\mathcal{I}(\mathcal{P}', \mathcal{L}) < C(m')^{11/15}n^{11/15}$. The base case of the induction is any m such that $n^{4/11} < m < n^{7/8}$, for which Lemma 2.1 gives

$$\mathcal{I}(\mathcal{P}, \mathcal{L}) \leq mn^{1/2} + n \leq 2m^{11/15}n^{11/15}.$$

We argue by contradiction; we will suppose that we have a configuration of $|\mathcal{P}|$ points $\mathcal{P}$ and $|\mathcal{L}|$ lines $\mathcal{L}$ satisfying $\mathcal{I}(\mathcal{P}, \mathcal{L}) = Cm^{11/15}n^{11/15}$. We then show, using the inductive hypothesis and the assumption $n^{7/8} < m < n^{8/7}$, that for a sufficiently large choice of C, independent of m and n, a contradiction occurs. We will work with explicit constants in the proof; we choose the constants for ease of comprehension, and we make no attempt to optimise them.

Suppose that $n^{7/8} < m < n^{8/7}$ and $I := \mathcal{I}(\mathcal{P}, \mathcal{L}) = Cm^{11/15}n^{11/15}$. Set $K := I/m$. We introduce two subsets of $\mathcal{P}$:

$$D := \{p \in \mathcal{P} : \text{ there are at most } 2^{-11}K \text{ lines through } p\}$$

and

$$E := \{p \in \mathcal{P} : \text{ there are at least } 2^{15}K \text{ lines through } p\}.$$

One can think of D as the set of points with a *dearth* of incidences, and E as the set of points with an *excess* of incidences.

It is evident that D contributes at most $2^{-11}Km = 2^{-11}I$ incidences to I. Similarly, we have the estimate $I \geq \mathcal{I}(E, \mathcal{L}) \geq 2^{15}K|E|$, which implies $|E| \leq 2^{-15}m$. By induction we have

$$\mathcal{I}(E, \mathcal{L}) < C\left(2^{-15}m\right)^{11/15} n^{11/15} < 2^{-11}I.$$

So E also contributes at most $2^{-11}I$ incidences to I.

Let $A := \mathcal{P}\backslash(E \cup D)$ be the remaining points. We define $c_1 = 2^{-11}$ and $c_2 = 2^{15}$. By definition of D and E, every point in A is incident to at least

c_1K and at most c_2K lines of $\mathcal{L}$. From the previous paragraph, we know that A contributes at least $\left(1 - 2 \cdot 2^{-11}\right) I$ incidences to I.

We repeatedly use Lemma 4.9to get the following sequence of grid-like subsets. Let $A_1 := A$. We iteratively choose $G_i \subset A_i$ as in Lemma 4.9, so there exist distinct points p_i, q_i such that G_i is covered by at most $2^{15}K$ lines from $\mathcal{L}$ through p_i, and by at most $2^{15}K$ lines from $\mathcal{L}$ through q_i. Then we set $A_{i+1} = A_i \backslash G_i$ and repeat. We terminate this process at the s-th step when $|A_{s+1}| \leq 2^{-15}m$ (allowing for the possibility that $s = 0$, which happens if $|A| \leq 2^{-15}m$, and the process is empty). This results in a sequence $A_1 \supseteq A_2 \supseteq \cdots \supseteq A_{s+1}$, with

$$|G_i| \geq \frac{c_1^4 K^4 |A_i|}{2^9 n^2} \geq \frac{(2^{-11})^4 K^4 (2^{-15}m)}{2^9 n^2} = \frac{K^4 m}{2^{68} n^2}$$

As the G_i are disjoint by construction, the process terminates after at most

$$s \leq \frac{m}{\min_i\{|G_i|\}} \leq \frac{2^{68} n^2}{K^4}$$

steps. It is a straightforward calculation to show that throughout the process, the conditions $K \geq 4n/(c_1|A_i|)$, $K \geq 8/c_1$ and $K^3 \geq 2^6 n^2/(c_1^3|A_i|)$ of Lemma 4.9 hold if C is chosen sufficiently large.

We may apply the inductive assumption to bound

$$\mathcal{I}(A_{s+1}, \mathcal{L}) < C(2^{-15}m)^{11/15} n^{11/15} = 2^{-11} I.$$

Thus the subsets $G_1, \ldots, G_s$ contribute at least $\left(1 - 3 \cdot 2^{-11}\right) I \geq I/2$ incidences, and in particular we have

$$I \leq 2 \sum_{i=1}^{s} \mathcal{I}(G_i, \mathcal{L}). \tag{4.4}$$

We now show that each G_i is projectively equivalent to a Cartesian product.

Recall from Chapter 2 that on the line at infinity ℓ_∞ in the projective plane $\mathbb{PF}^2$, there are two points α, β such that all lines through α (except for ℓ_∞) are horizontal lines in the affine plane, and the lines through β (except for ℓ_∞) are vertical lines in the affine plane.

For each i, we let τ_i be a projective transformation sending p_i and q_i to α and β.

The pre-image of the line at infinity is then the line $\overline{p_i q_i}$. From Lemma 4.9 we have $G_i \cap \overline{p_i q_i} = \emptyset$, so τ_i maps G_i into the affine plane. Also, if $\overline{p_i q_i}$

happens to be in $\mathcal{L}$, then it has no incidences with G_i, so we can ignore it when bounding $\mathcal{I}(G_i, \mathcal{L})$. The set $H_i = \tau(G_i) \subseteq \mathbb{F}^2$ is covered by $2^{15}K$ horizontal lines and $2^{15}K$ vertical lines, so it is contained in a Cartesian product $X_i \times Y_i$ with $|X_i| = |Y_i| \leq 2^{15}K$. Since projective transformations preserve incidences, we have $\mathcal{I}(H_i, \mathcal{L}) = \mathcal{I}(G_i, \mathcal{L})$.

We apply Theorem 4.5 to bound the incidences on each product $X_i \times Y_i$. In positive characteristic, the extra condition of Theorem 4.5 holds, since the assumption $m^{-2}n^{13} \ll p^{15}$ gives

$$|X_i||\mathcal{L}| \ll Kn \ll m^{-4/15}n^{26/15} \ll p^2.$$

Therefore, we can apply Theorem 4.5 to obtain (letting c^* denote the implicit constant in Theorem 4.5)

$$\mathcal{I}(G_i, \mathcal{L}) \leq \mathcal{I}(X_i \times Y_i, \mathcal{L}) \leq c^*(2^{15}K)^{3/4}(2^{15}K)^{1/2}n^{3/4} < c^*2^{20}K^{5/4}n^{3/4}.$$

Thus, using (4.4), we have (recalling that $K = I/m$ and that $s \leq 2^{68}n^2/K^4$)

$$I \leq 2\sum_{i=1}^{s} \mathcal{I}(G_i, \mathcal{L}) < 2 \cdot 2^{68}\frac{n^2}{K^4} \cdot c^*2^{20}K^{5/4}n^{3/4} = 2^{89}c^*\frac{m^{11/4}n^{11/4}}{I^{11/4}}.$$

Solving for I gives $I < C'm^{11/15}n^{11/15}$, for a constant C' that depends only on the constant c^* from Theorem 4.5, and not on C. Hence choosing $C > C'$ gives a contradiction to $I = Cm^{11/15}n^{11/15}$. This concludes the proof of Theorem 4.4.

4.8 Open questions

- An obvious open question is how to further improve the bounds in Theorem 4.4 and Theorem 4.5. We note that, even if the main term in the bound of Theorem 4.5 were improved to $O(|\mathcal{P}|^{2/3}|\mathcal{L}|^{2/3})$, our proof would not lead to the same bound in Theorem 4.4 (the result would be $O(|\mathcal{P}|^{8/11}|\mathcal{L}|^{8/11})$).

- Another interesting open problem, first posed by Bourgain [13], is whether similar bounds can be obtained for non-linear objects, like circles, conics, or other algebraic curves. Over $\mathbb{R}$ and $\mathbb{C}$ such bounds are known (see e.g. [24, 99]), and over $\mathbb{F}$ Bourgain [12] proved a qualitative theorem for hyperbolas. (For a very particular case, an explicit version was obtained by Shkredov [100], with applications to bounds on Kloosterman sums.)

- Yet another direction for research would be to impose more structure on the point set or line set, and to investigate how the maximum number of incidences can change accordingly. For example, if we have a point set where no s points are collinear, or a line set where no t are concurrent.

5

Applications of point-line incidences

5.1 Introduction

In this chapter we discuss the applications of the incidence bound in Chapter 4. Many of these applications and their proofs appeared in [108] and are elementary analogues of applications of the Szemerédi-Trotter bound in $\mathbb{R}$. The applications can be categorised into the following types:

1. Set expansion and sum-product estimates
2. Geometric applications
3. Applications to harmonic analysis and other areas of mathematics.

This thesis presents the opportunity to fully present the application of the general incidence bound Theorem 4.4 to harmonic analysis in finite fields (in the paper [108] we merely record the numerology of the restriction estimate).

We end this chapter with a review of subsequent work and important applications by other researchers.

5.2 Set expansion and sum-product estimates

The Sum-Product Phenomenon

The first application of the incidence bounds of the previous chapter is to the sum-product problem, and as such presents an opportunity in this thesis to discuss the history and context of this problem in the finite field setting. We will return to the sum-product problem in Chapter 7.

Sum-Product over $\mathbb{F}$

Recall from Chapter 3 the motto of the sum-product phenomenon is that *multiplication and addition cannot coexist*, and we expect this motto to carry over into all fields $\mathbb{F}$. In the setting of an arbitrary $\mathbb{F}$ are obstructions which do not occur in $\mathbb{R}$.

The first obstruction we encounter is unavoidable. Suppose we are in a finite field $\mathbb{F}$ and we take the (finite) set $A = \mathbb{F}$. Then $A + A,\ AA = \mathbb{F}$, and $|A + A|, |AA| = |A|$. Actually this obstruction is slightly more subtle: if $\mathbb{F}$ is any field and A a subfield of $\mathbb{F}$, then the numerology $|A + A| = |AA| = |A|$ is counter to the sum-product phenomenon that we might hope to expect. We will generally rule out this example by considering sets A which are bounded by the characteristic of the field, e.g. sets A of cardinality $|A| \leq p^{4/3}$, where p is the characteristic of $\mathbb{F}$.

Bourgain, Katz and Tao [17] showed that if this obstruction is not present (e.g. if A is a small enough subset of a prime residue field), then the sum set and product set cannot simultaneously be small. They also showed (Theorem 4.3) that the presence of subfields is the only obstruction towards a sum-product type phenomenon in a general finite field: as long as a set A is not a (translation or dilation of a) subfield, then it cannot be both 'strongly structured' in a multiplicative and additive sense.

Theorem 5.1 (Bourgain, Katz, Tao [17])**.** *Let p be an odd prime and let $A \subseteq \mathbb{F}_p$ satisfy $p^\delta < |A| < p^{1-\delta}$ for some $\delta > 0$.*

Then there exists $\epsilon = \epsilon(\delta) > 0$ and a constant $C = C(\delta) > 0$ such that

$$\max(|A + A|, |AA|) \geq C|A|^{1+\epsilon}.$$

This theorem was first quantified by Garaev [36], who proved that $\epsilon \geq \frac{1}{14}$. This was improved to $\frac{1}{13}$ by Katz and Shen [54], then to $\frac{1}{12}$ by Bourgain and Garaev [16]. In the previous record, a logarithmic term lingered, and Li [62] managed to achieve a bound of $\frac{1}{12}$ free from logarithmic factors. Rudnev [88] reduced the exponent further to $\frac{1}{11}$. At this stage, one might sensibly pause and try to guess the next exponent to be recorded. However, Roche-Newton, Rudnev and Shkredov [86] managed to defy the expected exponent by using Rudnev's point-plane theorem to show that $\epsilon \geq \frac{1}{5}$. We reproduce this bound using Theorem 4.5 using the argument of Elekes [25] sketched in the proof of Theorem 3.5.

We remark that a recent development of Rudnev, Shakan and Shkredov [91] improved the exponent $\frac{2}{10}$ to $\frac{2}{9}$.

Corollary 5.2. *Let $A \subset \mathbb{F}$ be a finite set. If $\mathbb{F}$ has positive characteristic p, assume that $|A| \ll p^{5/8}$. Then*

$$\max\{|A+A|, |AA|\}^2 \min\{|A+A|, |AA|\}^3 \gg |A|^6 .$$

In particular,

$$\max\{|A+A|, |AA|\} \gg |A|^{6/5}.$$

Moreover, if one of $|A+A|, |AA|$ is $O(|A|)$, then the other is $\Omega(|A|^{3/2})$.

Proof. We follow the proof as published in [108] and set

$$M_{\max} := \max\{|A+A|, |AA|\} \text{ and } M_{\min} := \min\{|A+A|, |AA|\} .$$

Define a point set and line set by

$$\mathcal{P} := (A+A) \times (A \cdot A) \quad \text{and} \quad \mathcal{L} := \{\ell_{ab} : (a,b) \in A \times A\},$$

where ℓ_{ab} is a line of the form $y = b(x-a)$.

In positive characteristic, we need to verify the condition $M_{min}|A|^2 \ll p^2$ of Theorem 4.5. If $M_{min} \gg |A|^{6/5}$ then we are done, so suppose that $M_{min} \ll |A|^{6/5}$. The required inequality $M_{min}|A|^2 \ll p^2$ follows from the assumption $|A| \ll p^{5/8}$.

The second condition of Theorem 4.5 is that $M_{min}M_{max}^2 \leq |\mathcal{L}|^3 = |A|^6$; if this fails, then $M_{max} \geq |A|^2$ and we are done.

The line $y = b(x-a)$ contains the point $(a'+a, ba')$ for any choice of $a' \in A$, so each of the $|A|^2$ lines gives at least $|A|$ incidences. Applying Theorem 4.5 gives

$$|A|^3 \leq I(\mathcal{P}, \mathcal{L}) \ll M_{min}^{3/4} M_{max}^{1/2} |A|^{6/4},$$

so

$$M_{min}^3 M_{max}^2 \gg |A|^6 . \tag{5.1}$$

□

The inequality $|A+A|^2|A \cdot A|^3 \gg |A|^6$ was obtained by Roche-Newton et al. [86] with the condition $|A| \ll p^{5/8}$, and $|A+A|^3|A \cdot A|^2 \gg |A|^6$ was obtained by Aksoy-Yazici et al. [118] with the condition $|A| \ll p^{3/5}$.

Equation (5.1) combines both these inequalities, and improves the condition for the second one. We note that the consistent exponent of $6/5$ across these three works is not entirely unexpected since they all rely on Rudnev's point-plane incidence bound.

Variations of the sum-product problem

A combination of both addition and multiplication of the set A should, by the philosophy of the sum-product conjecture, have any additive structure within the set (measured by looking at the sum-set) be destroyed by the multiplicative operation and vice versa. We will consider sets of the form $AA + A$ and $A(A + A)$.

Over arbitrary $\mathbb{F}$ we must take into consideration the same obstructions as before. The set A must be suitably small in terms of the characteristic of the field to avoid a saturation-type result, and to avoid the trivial lack of growth that arises if A is sub-field of $\mathbb{F}$.

This question was first considered by Barak, Impagliazzo and Wigderson [5], who used Konyagin's [59] (stronger version of Bourgain, Katz and Tao's original) sum-product estimate to prove that there exists $\epsilon > 0$ such that $|A + AA| \gg |A|^{1+\epsilon}$ for every $A \subset \mathbb{F}$ with $|A| < p^{0.99}$.

Barak, Impagliazzo and Wigderson were motivated by the application of extracting randomness from a small number of somewhat-random sources. Extractors are functions that take weak sources of randomness (i.e. random variables with low entropy[1], so heuristically 'far from random') and output a high source of randomness (i.e. the output is close to the uniform distribution). Explicitly, the application of the growth of $|AA + A|$ that Barak, Impagliazzo and Wigderson proved was to *dispersers*: a disperser is a function which takes in a constant number of inputs, and whenever the inputs are restricted to 'large-enough' sets, the image of the disperser is as large as possible [2].

Taking the function $f(x, y, z) = xy + z$ reveals the connection of dispersers to the growth of $AA+A$. Composing f with itself enough times, and identifying the field $\mathbb{F}$ with bit-strings $\{0, 1\}^n$ yields the desired disperser.

[1]The choice of entropy in this situation is *min-entropy*: the min-entropy of a random variable $X \in \{0, 1\}^n$ is $H_\infty(X) := -\min_{\omega \in \{0,1\}^n} \log_2(\mathbb{P}(X = \omega))$.

[2]A disperser with parameter set (k, l, m, n) is a function $D : \{0, 1\}^{nl} \to \{0, 1\}^m$ satisfying $D(A_1, \ldots, A_l) := \{D(a_1, \ldots, a_l) : a_i \in A_i, 1 \leq i \leq l\} = \{0, 1\}^n$ for all sets $A_1, \ldots, A_l \subseteq \{0, 1\}^n$ satisfying $|A_i| \geq 2^k$ for $i = 1, \ldots, l$.

We show now how incidence bounds are easily able to prove growth of the set $AA+A$ and its cousin $A(A+A)$. This refines the bounds of Roche-Newton, Rudnev and Shkredov [86], who proved the bound

$$|A+BC| \gg \min\{(|A||B||C|)^{1/2}, M^{-1}|A||B||C|, p\}$$

for $A, B, C \subset \mathbb{F}$, where $M = \max\{|A|, |B|, |C|\}$, and of Aksoy Yazici et al. [118], who proved the same bound for $A(B+C)$. We reprove both bounds, refining them by showing that the second term is not necessary as long as none of the sets is $\{0\}$.

We note that in finite fields this problem also has a 'Falconer-type' interpretation: how large must the set $A \subseteq \mathbb{F}_p$ be until A occupies (a positive proportion of) $\mathbb{F}_p$? Bienvenu et al. [10] show (amongst other results of a similar nature) that if $|A| = 0.3051p$, then $\mathbb{F}_p \setminus \{0\} \subseteq A(A+A)$.

Corollary 5.3. *Let $A, B, C \subset \mathbb{F}$ be finite sets, none of which equals $\{0\}$. If $\mathbb{F}$ has positive characteristic p, assume $|A||B||C| \ll p^2$. Then*

$$|A+BC| \gg (|A||B||C|)^{1/2} \quad \textit{and} \quad |A(B+C)| \gg (|A||B||C|)^{1/2}.$$

Proof. Without loss of generality, we assume that $|B| \geq |C|$ by interchanging B and C if necessary. Define a point set and line set by

$$\mathcal{P} := C \times (A+BC), \qquad \mathcal{L} := \{y = a + bx : (a,b) \in A \times B\}.$$

Each of the $|A||B|$ lines of $\mathcal{L}$ contains exactly $|C|$ points of $\mathcal{P}$, so there are $|A||B||C|$ incidences between $\mathcal{P}$ and $\mathcal{L}$.

In positive characteristic p, the condition $\min\{|C|, |A+BC|\} \cdot |\mathcal{L}| \ll p^2$ of Theorem 4.5 holds because of the assumption $|A||B||C| \ll p^2$. The other condition of Theorem 4.5 is that $|C||A+BC|^2 \leq (|A||B|)^3$, which we may assume, since otherwise we directly obtain $|A+BC|^2 > (|A||B|)^3|C|^{-1} \geq |A||B||C|$ using $|B| \geq |C|$. Thus we can apply Theorem 4.5 to get

$$|A||B||C| = \mathcal{I}(\mathcal{P}, \mathcal{L}) \ll |C|^{3/4}|A+BC|^{1/2}(|A||B|)^{3/4} + |A||B|.$$

If the first term dominates, rearranging gives the first inequality of the corollary. If the second term dominates, we have $|C| = O(1)$. Since $C \neq \{0\}$, we can pick a nonzero $c \in C$, and observe that $|A+cB| \geq \max\{|A|, |B|\} \gg (|A||B||C|)^{1/2}$. This finishes the proof of the first inequality.

The proof of the second inequality is similar. We first remove 0 from A, which does not affect the asymptotic behaviour (given that $A \neq \{0\}$). Then we define

$$\mathcal{P} := C \times (A(B+C)), \qquad \mathcal{L} := \{y = a(b+x) : (a,b) \in A \times B\},$$

noting that the lines are distinct because $0 \notin A$. As before, we then apply Theorem 4.5; checking the validity assumptions is almost identical to the previous application of the incidence bound, considering instead whether the inequality $|B||A(B+C)|^2 \leq (|A||B|)^3$ holds. □

We note that in the analogous question over $\mathbb{C}$, if we were to use the Szemerédi-Trotter theorem in the same manner, we would in fact achieve the same exponents, (see [111, Exercise 8.3.3]).

Over $\mathbb{R}$, stronger results are known; in the case where $A = B = C$, Roche-Newton and Warren [87] show that $|AA + A| \gg |A|^{3/2+1/194}$ whilst Murphy et al. [70] prove a stronger bound for the analogous problem: $|A(A+A)| \gg |A|^{3/2+1/186}$.

5.3 Geometric applications

Distinct distances and the pinned distance problem

The distinct distance problem in $\mathbb{R}^2$ was first asked by Erdős [29] in 1946: given a set of n points $\mathcal{P}$ in $\mathbb{R}^2$, how many distinct distances do they determine. We defer a discussion on the progress of this to Chapter 8, but mention only that Erdős conjectured that a lower bound for this problem should be $Cn(\log(n))^{-1/2}$ for an absolute constant $C > 0$, with the extremal bound attained by n points evenly distributed on a $\sqrt{n} \times \sqrt{n}$ integer lattice[3]. Since (an appropriate reinterpretation of) this example is valid also in finite field, it is thought that, under suitable restrictions, the same lower bound should hold over $\mathbb{F}_p$.

We write

$$d(q,r) = (q_x - r_x)^2 + (q_y - r_y)^2$$

for the squared Euclidean distance between two points $q = (q_x, q_y)$ and $r = (r_x, r_y)$, and

$$\Delta(\mathcal{P}) := |\{d(q,r) : q, r \in \mathcal{P}\}|$$

[3]We assume here that n is a square.

for the number of distances determined by $\mathcal{P}$. Guth and Katz [43] proved that for $\mathcal{P} \subset \mathbb{R}^2$ we have $\Delta(\mathcal{P}) \gg |\mathcal{P}|/\log|\mathcal{P}|$.

A related problem is the stronger 'pinned distance' problem, which asks for the existence of a point from which many distinct distances occur.

We write $\Delta_q(\mathcal{P}) = \{d(a,r) : r \in \mathcal{P}\}$ for the set of distances 'pinned' at a, and $\Delta_{\text{pin}} = \max|\Delta_a(\mathcal{P})|$, where the maximum is taken over all $a \in \mathcal{P}$.

The approach of [43] does not apply to this variant, and the best known bound is due to Katz and Tardos [55], who proved the estimate $\Delta_{\text{pin}}(\mathcal{P}) \gg |\mathcal{P}|^{0.86}$.

The finite field version of the pinned distance problem was first considered by Bourgain, Katz and Tao [17]. As with previous results, in the finite field analogue, we must restrict the size of the set of points to avoid the trivialisation of the problem that sub-fields present.

Bourgain, Katz and Tao proved a non-quantitative non-trivial bound: if $\mathcal{P} \subset \mathbb{F}_p^2$, with $|\mathcal{P}| = p^\alpha$ and $0 < \alpha < 2$, then $\Delta_{\text{pin}}(\mathcal{P}) \gg p^{\alpha(1/2+\epsilon)}$ for some $\epsilon = \epsilon(\alpha) > 0$. To avoid degeneracies arising from so-called 'isotropic points', Bourgain et al assumed that $p \equiv 3 \mod (4)$.

In the field $\mathbb{F} = \mathbb{F}_q$ where $q = p^r$ for some (large) prime p, if $|\mathcal{P}| \geq q^{4/3}$, then Hanson, Lund, and Roche-Newton [44] prove that $\Delta_{\text{pin}}(\mathcal{P}) \gg q$, which up to constants, is asymptotically optimal. We discuss this variant in further detail in the forthcoming Chapter 8.

For $\alpha < 4/3$ we present an explicit value for the ϵ in the statement of Bourgain, Katz and Tao [17].

Our proof is essentially that of [17], but we take some more care to deal with the case where -1 is a square in $\mathbb{F}$. When -1 is a square in $\mathbb{F}$, then it is possible for a set of points to determine exactly one distinct distance, namely the distance zero. Consider, for example, n points $\mathcal{P}$ in $\mathbb{C} \times \mathbb{C}$ all lying on the line $y = ix$ – let $\mathcal{P} = \{(a_1, ia_1), \ldots, (a_n, ia_n) : a_1, \ldots, a_n \in \mathbb{R}\}$. Then, for any $1 \leq j_1 < j_2 \leq n$, we calculate the distance between the pair of points indexed by j_1 and j_2: $d((a_{j_1}, ia_{j_1}), (a_{j_2}, ia_{j_2})) = (a_{j_1} - a_{j_2})^2 + (ia_{j_1} - ia_{j_2})^2 = 0$. This shows that the distance between any pair of points in the point set $\mathcal{P}$ is zero. This example is present in fields of characteristic $p \equiv 1 \mod (4)$, where i is defined to be a square root of -1.

Bourgain, Katz and Tao avoided this degeneracy by insisting that -1 is not a square[4].

[4]This is not stated in the journal version of [17], but it is mentioned in Section 7 of the later version arXiv:math/0301343v3.

We can rephrase the above obstruction by saying that all of the points lie on an *isotropic line*. The issue of isotropy is discussed in further detail in Section 8.7; for now, we shall define an isotropic line as a a line which satisfies the property that the distance between any pair of points lying on the line is zero.

Explicitly, for $r = (r_x, r_y) \in \mathbb{F}^2$,there are two isotropic lines passing through r,defined via the equations $(y - r_y) = \pm i \cdot (x - r_y)$, where $i^2 = -1$. We use λ_r and μ_r to denote the isotropic lines of r.

Corollary 5.4. *Let $\mathcal{P}$ be a set of points in $\mathbb{F}^2$. If $\mathbb{F}$ has positive characteristic, assume that $|\mathcal{P}| \ll p^{15/11}$. If -1 is a square in $\mathbb{F}$, assume that $\Delta(\mathcal{P}) \neq \{0\}$. Then*

$$\Delta_{pin}(\mathcal{P}) \gg |\mathcal{P}|^{8/15}.$$

Proof. For two distinct points $r = (r_1, r_2), s = (s_1, s_2) \in \mathbb{F}^2$, the field-analogue of the perpendicular bisector of r and s is the set

$$\ell_{rs} := \{q \in \mathbb{F}^2 : d(q, r) = d(q, s)\}.$$

The suggestive notation is testament to the fact that this set is a line: explicitly it is the line $x(2s_1 - 2r_1) + y(2s_2 - 2r_2) = s_1^2 + s_2^2 - r_1^2 - r_2^2$.

If -1 is not a square in $\mathbb{F}$, then for a fixed point r, distinct s give rise to distinct lines. When -1 is a square in $\mathbb{F}$, then for any two points s, t on one of the isotropic lines λ_r, μ_r we have $\ell_{rs} = \ell_{rt}$, which would cause a problem in the counting argument below.

We first deal with these isotropic lines. We know that, since the set of distances determined by $\mathcal{P}$ is not $\{0\}$, not all the points lie on an isotropic line. If any line (isotropic or regular) contains at least say $|\mathcal{P}|/3$ points of $\mathcal{P}$, and a point q outside that line, then $|\Delta_q(\mathcal{P})| \gg |\mathcal{P}|$. Combining these two observations, we can assume that no isotropic line contains more than $|\mathcal{P}|/3$ points of $\mathcal{P}$. Consequently, for any fixed $r \in \mathcal{P}$, there are at least $|\mathcal{P}|/3$ points not on λ_r or μ_r.

For a fixed $r \in \mathcal{P}$, consider the set of perpendicular bisectors determined by r and all non-isotropic vectors in $\mathcal{P}$:

$$\mathcal{L}_r := \{\ell_{rs} : s \in \mathcal{P}, d(r, s) \neq 0\}.$$

We have $|\mathcal{L}_r| \leq |\mathcal{P}| \ll p^{15/11}$. Applying Theorem 4.4 gives

$$\mathcal{I}(\mathcal{P}, \mathcal{L}_r) \ll |\mathcal{P}|^{22/15}.$$

Distinct points s with $d(r,s) \neq 0$ yield distinct lines in $\mathcal{L}_r$. Indeed, suppose (by translating the set of points if necessary) that $r = (0,0)$. Then ℓ_{rs} is given by $2s_1x + 2s_2y = s_1^2 + s_2^2$. Since s is non-isotropic (which is implied by $d(r,s) \neq 0$), we have that $s_1^2 + s_2^2 \neq 0$. From this calculation, it is clear that distinct s yield distinct lines.

An incidence $q \in \ell_{rs}$ corresponds to a semi-ordered triple of points $(q, \{r,s\})$ in $\mathcal{P}$ satisfying the relation $d(q,r) = d(q,s) \neq 0$. This is an analogue of vertices of an isosceles triangle (although we do not take care to exclude collinear points). Thus we have

$$\sum_{r\in\mathcal{P}} \mathcal{I}(\mathcal{P},\mathcal{L}_r) = \frac{1}{2}|\{(q,r,s) \in \mathcal{P}^3 : d(q,r) = d(q,s) \neq 0\}|$$
$$= \sum_{q\in\mathcal{P}} \frac{1}{2}|\{(r,s) \in \mathcal{P}^2 : d(q,r) = d(q,s) \neq 0\}|.$$

The factor $\frac{1}{2}$ in the above equation arises from counting ordered pairs (r,s) instead of sets $\{r,s\}$ in the algebraic interpretation of the incidence $q \in \ell_{rs}$.

For every $q \in \mathcal{P}$, the Cauchy-Schwarz inequality gives:

$$\left(\sum_{r\in\mathcal{P}} \sum_{\delta\in\Delta_q(\mathcal{P})\setminus\{0\}} \mathbb{1}_{d(q,r)=\delta}\right)^2 \leq |\Delta_q(\mathcal{P})||\{(r,s) \in \mathcal{P}^2 \colon d(q,r) = d(q,s) \neq 0\}|.$$

By the assumption that at least $|\mathcal{P}|/3$ points of $\mathcal{P}$ are not on an isotropic line through any given point in $\mathcal{P}$, for every $q \in \mathcal{P}$ we have that the left hand side is bounded below by $|\mathcal{P}|/3$.

Hence we may conclude that

$$|\mathcal{P}|^{37/15} \gg \sum_{r\in\mathcal{P}} \mathcal{I}(\mathcal{P},\mathcal{L}_r) \gg \sum_{q\in\mathcal{P}} \frac{|\mathcal{P}|^2}{|\Delta_q(\mathcal{P})|} \geq \frac{|\mathcal{P}|^3}{\max_q |\Delta_q(\mathcal{P})|}.$$

This gives $\Delta_{\text{pin}}(\mathcal{P}) \gg |\mathcal{P}|^{8/15}$. □

One might naturally ask whether this proof is strengthened in the case where the point set is a Cartesian product. We can perform the same proof as above, yielding $\Delta_{\text{pin}}(A \times A) \gg |A|^{5/4}$ under the condition $|A| \ll p^{2/3}$; this is a strengthening of a result of Aksoy Yazici et al. [118, Corollary 13(a)]. However, Petridis [81] improved this to $\Delta_{\text{pin}}(A \times A) \gg |A|^{3/2}$, using Rudnev's point-plane theorem directly. His proof does not seem to apply to unstructured point sets as in Corollary 5.4.

Corollary 5.4 is improved in Chapter 8.

Beck's theorem

Beck's Theorem of Two Extremes [7] (referred to in this work as 'Beck's Theorem') is a structural statement about points in the real plane. Beck proved that either $\Omega(|\mathcal{P}|)$ points are on a line, or else $\mathcal{P}$ determines $\gg |\mathcal{P}|^2$ lines. We say that a line ℓ is determined by $\mathcal{P}$ if there are at least two points of $\mathcal{P}$ lying on ℓ. (Any two points determine a line, and a line is determined by any pair of points lying on it.)

Beck's original proof was proved independently of the Szemerédi-Trotter theorem; in fact both papers were published in the same 1983 issue of Combinatorica. However, the Szemerédi-Trotter theorem can also be used to prove Beck's theorem. We will prove an analogue of Beck's theorem over arbitrary fields using this method.

Beck's theorem in finite fields has previously been investigated: in $\mathbb{F}_p^2$, Helfgott and Rudnev [46] established that if $|A| < \sqrt{p}$ and $\mathcal{P} = A \times A$ (so no line has $\Omega(|\mathcal{P}|)$ points), then $\mathcal{P}$ determines $\Omega(|\mathcal{P}|^{1+1/267})$ lines. Jones [51] strengthened the exponent and removed the Cartesian product condition, proving that either $\mathcal{P}$ has $\Omega(|\mathcal{P}|)$ points on a line, or $\mathcal{P}$ determines $\Omega(|\mathcal{P}|^{1+1/109})$ lines. Jones used the approach using point-line incidences; since Jones' incidence bound Theorem 4.2 is improved using [86, Theorem 6], his bound on Beck's theorem is automatically updated to an exponent of $1 + 1/53$.

We note that for large point sets the issue of an analogue of Beck's theorem is resolved: Alon [2] proved that any point set $\mathcal{P} \subset \mathbb{F}_q^2$ of size $|\mathcal{P}| > q$ determines $c|\mathcal{P}|^2$ lines, with c depending on $|\mathcal{P}|/q$.

We present first a general result for an unstructured point set, improving on [51]. We will use the incidence bound Theorem 4.4.

Although the proof works for $|\mathcal{P}| \ll p^{7/6}$, it is weaker than the result of Alon [2] for $|\mathcal{P}| > p$.

Corollary 5.5. *Let $\mathcal{P}$ be a set of m points in $\mathbb{F}^2$. If $\mathbb{F}$ has positive characteristic p, suppose that $m \ll p^{7/6}$. Then one of the following is true:*

(i) $\mathcal{P}$ has $\Omega(m)$ points on a line;

(ii) $\mathcal{P}$ determines $\Omega(m^{8/7})$ lines.

Proof. Let $|\mathcal{P}| = m$, and let $\mathcal{L}$ be the set of lines determined by $\mathcal{P}$. We partition the set of lines in $\mathcal{L}$ into $\lfloor \log_2 m \rfloor$ sets $\mathcal{L}_j \subseteq \mathcal{L}$ according to the number of

points on each line:

$$\mathcal{L}_j = \{\ell \in \mathcal{L} \colon 2^j \leq |\ell \cap \mathcal{P}| < 2^{j+1}\}\,.$$

Consider the popular line sets $\mathcal{L}_j$ with $|\mathcal{L}_j| > m^{7/8}$. We can assume that each $|\mathcal{L}_j| < m^{8/7}$, since otherwise (ii) holds. Thus, in positive characteristic, we can use the assumption $m \ll p^{7/6}$ to get $m^{-2}|\mathcal{L}_j|^{13} \ll p^{15}$. We are now able to apply Theorem 4.4 to $\mathcal{P}$ and $\mathcal{L}_j$. Every line in $\mathcal{L}_j$ has at least 2^j incidences with $\mathcal{P}$, and so Theorem 4.4 gives

$$2^j|\mathcal{L}_j| \ll m^{11/15}|\mathcal{L}_j|^{11/15}\,.$$

Rearranging gives

$$|\mathcal{L}_j| \ll \frac{m^{11/4}}{(2^j)^{15/4}}\,.$$

Each line in $\mathcal{L}_j$ contains $\sim 2^{2j}$ pairs of points; all the lines in $\mathcal{L}_j$ together contain $O(m^{11/4}(2^j)^{-7/4})$ pairs of points of $\mathcal{P}$.

For $\mathcal{L}_j$ with $m^{1/2} \leq |\mathcal{L}_j| \leq m^{7/8}$, the trivial incidence bound Lemma 2.1 gives $2^j|\mathcal{L}_j| \ll m^{1/2}|\mathcal{L}_j|$, so $2^j \ll m^{1/2}$. Hence, the number of pairs of points on lines of $\mathcal{L}_j$ is $O(m^{15/8})$.

If $|\mathcal{L}_j| \leq m^{1/2}$, then we can again use the trivial incidence bound to get $2^j|\mathcal{L}_j| \ll m$, so $|\mathcal{L}_j| \ll 2^{-j}m$. This implies that the number of pairs of points on lines of $\mathcal{L}_j$ is $O(2^j m)$.

Now, let C be a large constant and let U be the union of all $\mathcal{L}_j$ satisfying $Cm^{3/7} \leq 2^j \leq m/C$. By the estimates above, the number of pairs of points on lines contained in U is at most

$$O\left(\frac{m^{11/4}}{(Cm^{3/7})^{7/4}} + m^{15/8}\log m + \frac{m}{C}\cdot m\right) = O\left(\frac{m^2}{C}\right)\,.$$

For sufficiently large C, this quantity is less than $\frac{1}{2}\binom{m}{2}$.

Thus, the remaining $\Omega(m^2)$ pairs of distinct points of $\mathcal{P}$ must lie outside U. Then one of the following two things must happen: either a positive proportion of these pairs are supported on lines containing more than m/C points, or a positive proportion of pairs lie on lines with less than $Cm^{3/7}$ (and at least two) points.

In the first case, there is a line containing at least m/C points, and (i) holds. In the second case, the number of distinct lines determined by pairs of points of $\mathcal{P}$ is

$$\Omega\left(\frac{m^2}{(Cm^{3/7})^2}\right) = \Omega\left(m^{8/7}\right)$$

and so (ii) holds.

This completes the proof of the corollary. □

A natural question is to ask whether one can gain a stronger result using the stronger Theorem 4.5 instead of Theorem 4.4. To be in this situation, we need the point set to have a Cartesian-product structure. In the case that $\mathcal{P} = A \times A$, our proof of Corollary 5.5 gives the same result as in [118]; however we can obtain a more precise result when $\mathcal{P}$ is an asymmetric Cartesian product of the form $\mathcal{P} = A \times B$.

Using Rudnev's point-plane incidence bound, Aksoy Yazici et al. [118] were able to show that $\mathcal{P} = A \times A \subseteq \mathbb{F}^2$ determines $\Omega(|\mathcal{P}|^{3/2})$ lines over an arbitrary field $\mathbb{F}$ (as long as $|\mathcal{P}| \ll p^{4/3}$ in positive characteristic). (Note that the Cartesian structure of the point set prevents $\gg |\mathcal{P}|$ points being collinear.) Using the same proof as in Corollary 5.5 we show the following Beck-type result.

Corollary 5.6. *Let $\mathcal{P} = A \times B$ be a set of points in $\mathbb{F}^2$ where $|A| \leq |B|$ and $|A|^2|B| \ll p^2$ in positive characteristic p. Then one of the following must happen:*

(i) *$\mathcal{P}$ has $\Omega(|A|^{5/6}|B|^{2/3})$ points on a line;*

(ii) *$\mathcal{P}$ determines $\Omega(|A||B|^2)$ lines.*

We remark that $\mathcal{P}$ cannot have more than $|B|$ points on a line, so the first condition is vacuous, unless $|A| \leq |B|^{2/5}$.

Proof. Using the same $\mathcal{L}_j$ as defined in the proof of Corollary 5.5, we find that either $|\mathcal{L}_j| \ll |A|^{1/3}|B|^{2/3}$, or else we are within the range of applicability of Theorem 4.5, and so

$$|\mathcal{L}_j| \ll \frac{|A|^3|B|^2}{2^{4j}}\,.$$

The assumption $|A|^2|B| \ll p^2$ enables the verification of the condition of Theorem 4.5 than $|A|\#\{\text{lines}\} \ll p^2$. To see this, we use the trivial incidence bound exactly as in Corollary 5.5.

As before, we take U to be a union of lines in $\mathcal{L}_j$. This time, we parameterise the set U differently, allowing sets $\mathcal{L}_j$ satisfying $C\mu \leq 2^j \leq M/C$; we will chose μ, M later.

The number of pairs of points on lines contained in U is at most

$$O\left(\frac{|A|^3|B|^2}{C^2\mu^2} + \frac{M^2|A|^{1/3}|B|^{2/3}}{C^2}\right).$$

We choose $\mu^2 = |A|$ and $M^2 = |A|^{5/3}|B|^{4/3}$; for sufficiently large C, the above quantity is less than $\frac{1}{2}\binom{|A||B|}{2}$, the total possible number of pairs of points.

Hence there is either a line with at least $|A|^{5/6}|B|^{2/3}$ points, or the number of distinct lines is $\gg |A||B|^2$. □

5.4 Collinear Quadruples

We conclude this section with a theorem of Petridis [79, 70] regarding collinear quadruples in $A \times A$. Theorem 4.5 yields a bound of $O(|\mathcal{P}|^{5/2}k^{-4} + |\mathcal{P}|k^{-1})$ for the number of k-rich lines with respect to the Cartesian Product point set $\mathcal{P} = A \times A$.

Petridis was the first to realise that Theorem 4.5 was best-suited to 'fourth moment' applications, via an optimal bound on the number of collinear quadruples in $A \times A \subseteq \mathbb{F}^2$, under suitable restrictions on $|A|$ in terms of the characteristic of the field $\mathbb{F}$. He showed that the number of collinear quadruples in $A \times A$ is bounded by $O(|A|^8p^{-2} + |A|^5\log(|A|))$. That the number of collinear quadruples is estimated by a 'fourth moment' sum is illustrated by the forthcoming equation 5.2. For comparison, the Szemerédi-Trotter theorem is best-suited to 'third moment' applications, as demonstrated by the sharp bound on collinear quadruples derived from it in Theorem 2.12.

We demonstrate the role of incidences to the problem of bounding the number of collinear quadruples in a point set by considering the analogous asymmetric problem $\mathcal{P} = A \times B$.

Theorem 5.7. *Suppose that $A, B \subseteq \mathbb{F}$ are finite sets with $|A| \leq |B|$ satisfying the additional constraint $|A|^3|B|^2 \leq p^2$ in positive characteristic.*

Then the number of collinear quadruples in $A \times B$ is $O(|A||B|^4 + |A|^3|B|^2\log|B|)$.

Proof. We first count the $O(|A||B|^4)$ collinear quadruples arising from horizontal and vertical lines and from quadruples in which not all points are distinct.

Let $\mathcal{L} = \mathcal{L}(\mathcal{P})$ be the set of lines determined by $\mathcal{P} = A \times B$, and for $i = 1, \ldots, \lceil \log_2 |B| \rceil$, define the line sets:

$$\mathcal{L}_i := \{\ell \in \mathcal{L} : 2^i \leq |\ell \cap \mathcal{P}| < 2^{2i}\}.$$

The number of collinear quadruples in $\mathcal{P}$ is

$$Q(A \times B) := \sum_{\ell \in \mathcal{L}} \binom{|\mathcal{P} \cap \ell|}{4} < \sum_{i=1}^{\lceil \log_2(N) \rceil} 2^{4i} |\mathcal{L}_i| \,. \tag{5.2}$$

We claim that the conditions of Theorem 4.5 are satisfied for each i. For this to hold, we require that $|A||\mathcal{L}_i| \ll p^2$. This condition follows from the trivial incidence bound and the hypotheses: Lemma 2.1 yields the inequality $|\mathcal{L}_i| \leq |A|^2|B|^2 2^{-2i}$, and so $|A||\mathcal{L}_i| \leq |A|^3|B|^2 \ll p^2$.

Applying Theorem 4.5, we see that $|\mathcal{L}_i| \ll |A|^3|B|^2 2^{-4i}$, and so

$$Q(A \times B) \ll \sum_{i=1}^{\lceil \log_2 |B| \rceil} 2^{4i} \frac{|A|^3|B|^2}{2^{4i}} \leq |A|^3|B|^2 \log(|B|) \,.$$

□

5.5 Open Questions

- The most major open question related to this chapter is the sum-product problem over arbitrary fields and the correct bounds on related questions.
- A problem which has received less attention in the literature is a Beck-type bound in positive characteristic. The bound we prove is likely far from optimal. This geometric question may well be resolvable with modern machinery using tools from algebraic geometry.
- Incidence bounds have been used by the computer science community to prove bounds relating to pseudorandomness. Despite the much stronger bounds of Theorems 4.4 and 4.5, for a computer science application we would require strong bounds when the characteristic of the field is 2. It would be interesting to overcome this obstacle.

6

Energy decomposition of a set

6.1 Introduction

The sum-product phenomenon has already been amply mentioned throughout this work, and uses the cardinality of the sum and product sets to quantify this question. Recall that the sum-product conjecture expects one of the sum set or the product set to be large in order to evidence the notion that multiplication and addition cannot coexist in a set.

We have seen that the additive (resp. multiplicative) energy of a set is a finer measure of the additive (resp. multiplicative) structure of a set. We recall the definition of the additive energy (Definition 3.7) and allow the reader's imagination to extract the multiplicative energy:

$$\mathsf{E}^+(A,B) := \sum_{x\in A+B} r^2_{A+B}(x) = \sum_{x\in A-B} r^2_{A-B}(x) = \sum_x r_{A-A}(x) r_{B-B}(x)$$

A theme of this chapter will be that *a set with large energy has structure*. Our aim in this chapter is to study an energy-formulation of the sum-product conjecture: to what extent can addition and multiplication coexist when quantified by the energy of a set?

In Example 3.10, we studied a set A which was the union of an equal-sized arithmetic and geometric progression. This set clearly contains both additive and multiplicative structure, which we capture by having large multiplicative and additive energy estimates. Yet the sum-product phenomenon conjectures

the impossibility of this coexistence! At this point, one might dismiss the possibility of an energy formulation of a sum-product phenomenon.

However, a recent result of Balog and Wooley [4] should dissuade the reader of the hopelessness of this task. Balog and Wooley show that, if one is willing to pass to a large subset of the original set, then one of the multiplicative or additive energies of this subset must be small.

Theorem 6.1 (Balog, Wooley [4])**.** *Let $A \subseteq \mathbb{R}$ and $\delta = 2/33$. Then there exists $A' \subseteq A$ such that $|A'| \geq |A|/2$ and*

$$\min(\mathsf{E}^+(A'), \mathsf{E}^\times(A')) \lesssim |A|^{3-\delta}\,.$$

Improving this theorem is the main work of this chapter. This chapter is joint work with Misha Rudnev and Ilya Shkredov and parts of the proofs appear in the publication [92].

Structure of this chapter

We begin with a discussion on the energy formulation of the sum-product conjecture and state the main result of this chapter. The key tool we use is a decomposition result; we will apply an iterative argument to this decomposition result to prove the main energy estimate. Section 6.3 is devoted to the proof of a new decomposition result.

In Chapter 7, we will consider applications of a result of this type: to expansion and to the sum-product phenomenon.

6.2 On the energy formulation and Balog–Wooley decomposition

To better understand the statement of Theorem 6.1, let us return once more to the example of a set A that is a union of an arithmetic and geometric progression of equal size. For $N \geq 3$, let

$$A = \{1, 2, \ldots, N\} \cup \{2^N, 2^{N+1}, \ldots, 2^{2N}\}$$

Let $A' = \{1, 2, \ldots, N\}$; then by Solymosi's [104] theorem, in particular equation (3.5) of Section 3.3, we have $\mathsf{E}^\times(A') \leq 8\log(|A'|)|A'|^2 \lesssim |A|^2$. (Note that in this example, $\mathsf{E}^+(A')$ is maximal.)

This example is suggestive of the fact that Balog and Wooley's theorem is non-optimal. In fact, it is not at all obvious what the optimal value of δ should be. To consider this, we begin by examining an instructive construction of Balog and Wooley [4].

Example 6.2 (Balog–Wooley [4])**.** *Let $I = \{n^2, n^2+1, \dots 2n^2-1\}$ and let A be the union of n disjoint dilates of I by $1, 2, \dots, 2^{n-1}$:*

$$A = I \cup 2I \cup \cdots \cup 2^{n-1}I\,.$$

Figure 6.1: The Balog Wooley example: illustration of $A \times A$ with $n = 6$

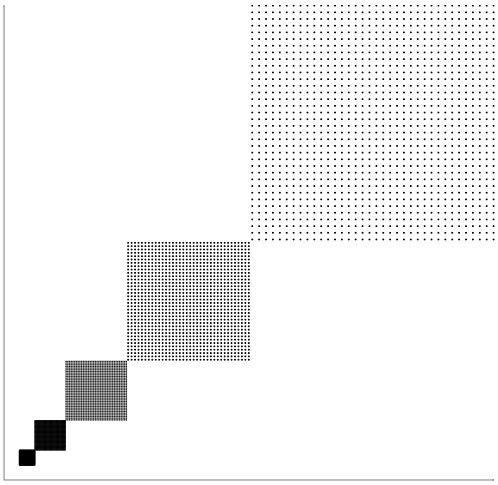


Figure 6.2 illustrates the structure of the set $A \times A$. The set A is simultaneously highly multiplicative and highly additive, in a stronger sense than our previous example of the union of an arithmetic and geometric progression.

For any set $B \subseteq A$ satisfying $|B| \geq |A|/2$ we claim that

$$\mathsf{E}^+(B) \gg |A|^{7/3} \qquad \text{and} \qquad \mathsf{E}^\times(B) \gtrsim |A|^{7/3}\,.$$

Indeed, first note that $|A| = n^3$, and let $A_j := 2^j I$. Then there exists an indexing set $J \subseteq \{0, \dots, n-1\}$ of size at least $n/2$ so that for each $j \in J$, $|B \cap A_j| \geq n^2/2$. Then, using the Cauchy-Schwarz inequality Lemma 3.11 in the second inequality, we have

$$\mathsf{E}^+(B) \geq \sum_{j \in J} \mathsf{E}^+(B \cap A_j) \geq \sum_{j \in J} \frac{|B \cap A_j|^4}{|(B \cap A_j) + (B \cap A_j)|} \gg n^7\,.$$

In fact, the choice of constant $1/2$ in the size of J is inconsequential; a different choice of constant would yield a different bound in the $\gg$ notation.

To estimate $\mathsf{E}^\times(B)$*, we use the the Cauchy-Schwarz inequality Lemma 3.11 and the bound* $|AA| \sim n^5$ *to obtain*

$$\mathsf{E}^\times(B) \geq \frac{|B|^4}{|BB|} \gg \frac{n^{12}}{|AA|} \sim n^7 .$$

An explicit bound for the logarithmic factor hidden in this notation is given by Ford [33] – see Example 3.2.

This example leads to the following conjecture by Rudnev, Shkredov and the author in [92]:

Conjecture 6.3. *For* $A \subseteq \mathbb{R}$ *there exists a subset* $A' \subseteq A$ *such that* $|A'| \geq |A|/2$ *and*

$$\min(\mathsf{E}^+(A'), \mathsf{E}^\times(A')) \lesssim |A|^{7/3} .$$

Balog–Wooley decomposition

Balog and Wooley formulated their result as a decomposition result, from which Theorem 6.1 is an immediate consequence. They show that any finite set $A \subseteq \mathbb{R}$ admits a decomposition into two parts[1] $B \sqcup C$ so that one of B or C is large *and* lacking in one of multiplicative or additive structure.

Theorem 6.4 (Balog–Wooley decomposition [4])**.** *Let* $A \subset \mathbb{R}$ *be a finite set and* $\delta = 2/33$*. Then there are two disjoint subsets* B *and* C *of* A *such that* $A = B \sqcup C$ *and*

$$\max\{\mathsf{E}^+(B), \mathsf{E}^\times(C)\} \ll |A|^{3-\delta}(\log|A|)^{1-\delta}$$

and

$$\max\{\mathsf{E}^+(B,C), \mathsf{E}^\times(B,C)\} \ll |A|^{3-\delta/2}(\log|A|)^{(1-\delta)/2} .$$

Note that one could take $|B| \ll 1$, if, say, A is a geometric progression. Most importantly, this theorem says that at least half of the set A is either highly non-additive or highly non-multiplicative.

The proof of Theorem 6.4 uses the Balog–Szemerédi–Gowers Theorem [3, 38, 95] described in Section 3.3. In [58], Konyagin and Shkredov applied a different method and were able to obtain an improvement; they showed that δ could be taken to be at least $1/5$.

[1] We use the notation $A = B \sqcup C$ to denote that A is a disjoint union of B and C.

Main Decomposition Results

In [92], we further improve the value of δ.

Theorem 6.5. *Let $A \subset \mathbb{C}$ be a finite set, and $\delta = 1/4$. Then there are two disjoint subsets B and C of A such that $A = B \sqcup C$ and*

$$\max\{\mathsf{E}^+(B), \mathsf{E}^\times(C)\} \lesssim |A|^{3-\delta}$$

We also prove an analogous result over arbitrary fields.

Theorem 6.6. *Let $A \subset \mathbb{F}$ be a set and $\delta = 1/5$. In positive characteristic p, assume in addition that $|A| \leq p^{5/8}$.*

Then there are two disjoint subsets B and C of A such that $A = B \sqcup C$ and

$$\max\{\mathsf{E}^+(B), \mathsf{E}^\times(C)\} \lesssim |A|^{3-\delta}\,.$$

Both of these theorems are proved in the same manner; the improvement over $\mathbb{R}$ is a consequence of using the Szemerédi-Trotter incidence theorem in place of Rudnev's points/planes theorem. The proofs rely on an intermediate result which guarantees the existence of a 'large' set $A_1 \subseteq A$ which has 'small' energy. The largeness of the set A_1 is controlled by the energy of A; if A_1 is quantitatively large with respect to the multiplicative energy of A, then the subsequent additive energy of A_1 will be small. Here, multiplication and addition may be swapped. This concept is the content of the forthcoming Propositions 6.8 and 6.9.

It is relatively simple to derive similar types of result, and we demonstrate one such example.

Corollary 6.7. *Let $A \subseteq \mathbb{F}$ be a finite set with, in positive characteristic, the additional constraint that $|A| \leq p^{3/5}$.*

(i) Then there exists a partition of A into $B \sqcup C$, where $|B|, |C| \geq |A|/3$ and

$$\mathsf{E}^+(B)\ \mathsf{E}^\times(C)^{3/2} \lesssim |A|^7\,. \tag{6.1}$$

Addition and multiplication may be swapped in the above (for different B, C).

(ii) There exists a (possibly different) partition of A into $B' \sqcup C'$, where $|B'|, |C'| \geq |A|/3$ and

$$\mathsf{E}^+(B')\ \mathsf{E}^\times(C') \lesssim |A|^{28/5}\,. \tag{6.2}$$

When $\mathbb{F} = \mathbb{C}$, the exponent $28/5$ can be replaced by $11/2$.

6.3 Proof of Decomposition Results

Initial Decomposition

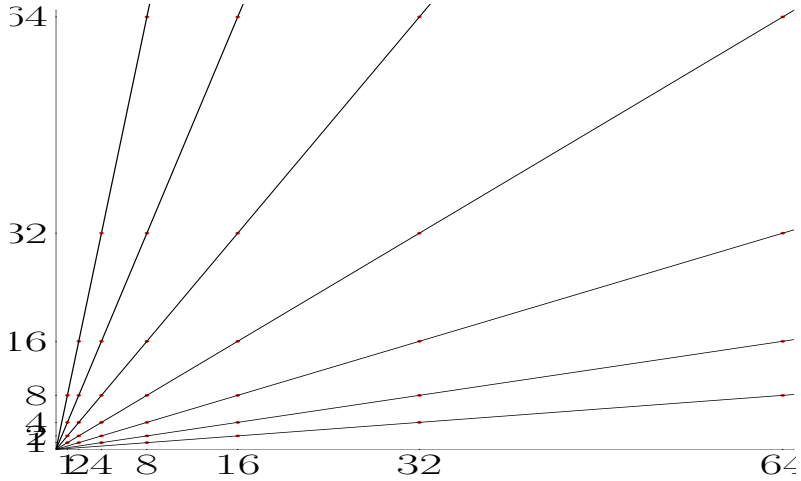
As previously discussed, the key concept in the proof of Theorems 6.5 and 6.6 is a technique to extract a large subset whose structure and size is controlled by an interplay of additive and multiplicative energy. The arbitrary field version of this idea is somewhat more straightforward, so we will prove this (Proposition 6.8) and then indicate the changes needed for the improvement in $\mathbb{R}$, Proposition 6.9.

Proposition 6.8. *Let $A \subset \mathbb{F}$ be a finite set. If the characteristic of $\mathbb{F}$ is $p > 0$, suppose in addition that $|A|^6 \leq p^2\ \mathsf{E}^\times(A)$.*
Then there is $A_1 \subseteq A$ such that $|A_1|^2 \gtrsim \mathsf{E}^\times(A)|A|^{-1}$ and

$$\mathsf{E}^+(A_1) \lesssim |A_1|^{11/2}|A|^{3/2}(\mathsf{E}^\times(A))^{-3/2}\,. \tag{6.3}$$

Proof. We begin by describing the procedure that will return the desired set A_1. To enhance comprehension and hopefully clarity, this description is accompanied by diagrams.

Figure 6.2: Taking $A = \{1, 2, 4, 8, 16, 32, 64\}$, we restrict to points in $A \times A$ supported on slopes containing a uniform number of points. In this example, each slope contains between $t + 1$ and $2t$ points of $A \times A$, with $t = 4$.

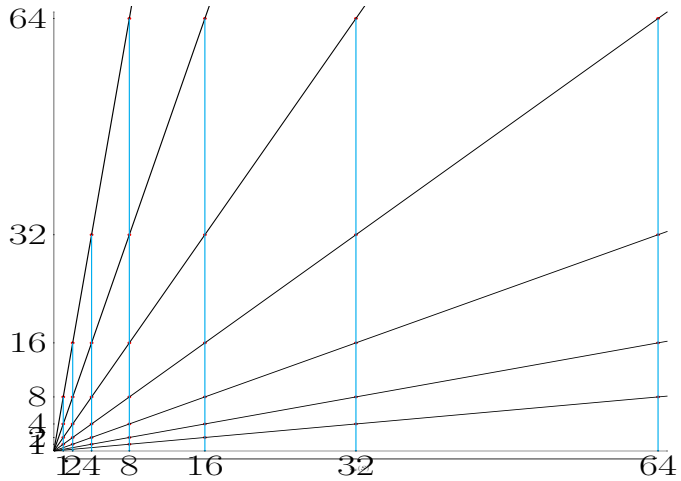


First, using the dyadic pigeonhole principle, we choose a set $P \subseteq A/A$ which supports a $(\gtrsim 1)$ proportion of $\mathsf{E}^{\times}(A)$, with each $p \in P$ having approximately t realisations from A/A for some integer $t \geq 1$.
Indeed,

$$\max_i\{2^{2i}|P_i|\} \leq \mathsf{E}^{\times}(A) \leq \sum_{i=0}^{\lfloor \log |A| \rfloor} 2^{2i+2}|P_i| \leq 2\log(|A|)\max_i\{2^{2i}|P_i|\}\,,$$

where $P_i = \{\rho \in A/A : 2^i < r_{A/A}(\rho) \leq 2^{i+1}\}$. We set P and t to be the P_i and the 2^i of this maximal case.

Figure 6.3: Project the remaining points to the x-axis



Let $S \subseteq A \times A$ be the set of points supported on lines passing through the origin with slopes in P; we have $|P|t \leq |S| < 2|P|t$.
Let $\pi_x : S \mapsto A$ be the projection of points of S to the x-axis: $\pi_x(s_x, s_y) = s_x$. The projection π_y is similarly defined as the projection to the set of ordinates.

Consider the set $A_x = \pi_x(S)$ of abscissae of S. We perform another dyadic pigeonholing argument to find a set $A' \subseteq A_x$ and a number q' such that $q' < |A \cap xP| \leq 2q'$ for all $x \in A'$.

That is, let $(A_x)_j := \{a_x \in A_x : 2^j < |A \cap a_x P| \leq 2^{j+1}\}$. Then $\sum_j 2^j |(A_x)_j| < |S| \leq \sum_j 2^{j+1}|(A_x)_j|$, where the indexing set j runs over $j = -1, \ldots \lceil \log(|A|) \rceil$. Thus, there exists an index j_0 so that $|S| \gtrsim 2^{j_0}|(A_x)_{j_0}|$. We set $q' : 2^{j_0}$ and $A' = (A_x)_{j_0}$.

The set A' is the set of popular abscissae of S – the vertical line through each $x \in A'$ contains $\sim q'$ points of S. Note that $|A'|q' \sim |S|$.

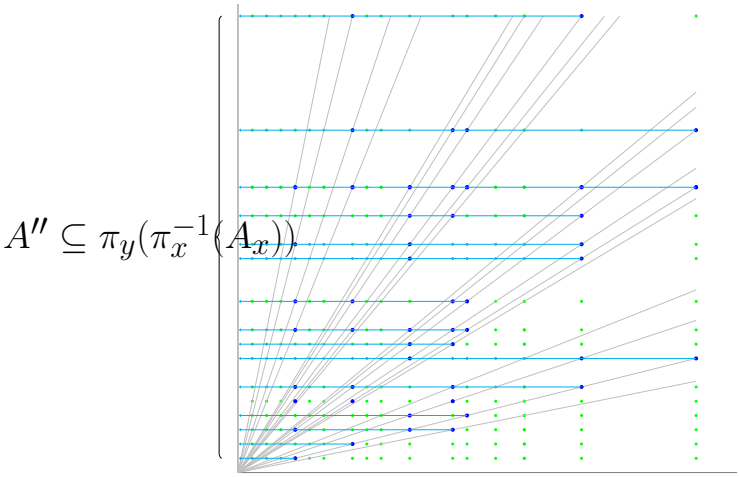
Observe that $q' \leq m := \min\{|A|, |P|\}$. If $q' \leq |A'|$ then we take $A_1 = A'$ and set $q = q'$.

It may be the case however that $q' > |A'|$. In this scenario we have a somewhat (narrow) rectangular set of points, where the height of this rectangle is greater than its width. If $q' > |A'|$ then we once again apply the dyadic pigeonhole principle to regularise our set. Let

$$S_x = \pi_x^{-1}(A') \cap S$$

be the set of points in S and abscissae in A' (so S_x is supported on on $|A'| < q'$ vertical lines). Consider now the ordinates of this set; by the dyadic pigeonhole principle, there exists an integer q'' and $A'' \subseteq \pi_y(S_x)$. That is, the horizontal line through each $a'' \in A''$ contains $\sim q''$ points of S_x. Crucially, $|A''|q'' \sim |S_x|$.

Figure 6.4: Project the points on the rectangular grid horizontally and then use the dyadic pigeonhole principle to obtain A''



We have

$$q'|A'| \sim q''|A''|\,,$$

and since $q' > |A'|$, we have that $|A''| \gtrsim q' > |A'|$, and so $q'' \lesssim q' < |A''|$. We then take $A_1 = A''$ and $q = q''$, concluding that $|A_1| \gtrsim q$, and thus

$$|A_1|^2 \gtrsim q|A_1| \sim |S| \geq |P|t = \frac{|P|t^2}{t} \sim \frac{\mathsf{E}^{\times}(A)}{t} \geq \frac{\mathsf{E}^{\times}(A)}{|A|}\,. \tag{6.4}$$

At this point we pause to observe that we have a set A_1 of the desired cardinality; it remains to prove that A_1 also has the sought additive structure.

We have by construction that each member of A_1 can be represented at least q times as an element of A/P (assuming that is, that we chose A_1 to be the set of popular ordinates; had we chosen A_1 to be the set of popular abscissae, then we would have that each member of A_1 is representable at least q times as an element of the set AP; note that $P = P^{-1}$ and so we do not need to concern ourselves with this technicality).

Thus we can write

$$\begin{aligned}\mathsf{E}^{+}(A_1) &= |\{(a,a',b,b') \in A_1^4 : a+b = a'+b'\}| \\ &\leq q^{-2}|\{(a,a',r,r',\alpha,\alpha') \in A_1^2 \times P^2 \times A^2 : a + r\alpha = a' + r'\alpha'\}| \\ &= q^{-2}\mathcal{I}(|A_1||P||A| \text{ points }, |A_1||P||A| \text{ planes })\,,\end{aligned}$$

where the planes in question have equation $a + rx = y + \alpha' z$ and points $(x,y,z) \in A \times A_1 \times P$.

We will apply Rudnev's incidence bound, Theorem 2.9. In positive characteristic we have the further requirement of $|A||A_1||P| < p^2$, i.e. the number of points is at most p^2. We justify this now.

We have that

$$|A||P||A_1| < |A|^2|P| \lesssim |A_1|^4|A|^2\mathsf{E}^{\times}(A)^{-1} \leq |A|^6\mathsf{E}^{\times}(A)^{-1}\,,$$

using the trivial inequality $|A_1| \leq |A|$ and the observation that

$$|P| = \frac{|P|^2t^2}{|P|t^2} \sim \frac{|S|^2}{\mathsf{E}^{\times}(A)} \leq \frac{|A_1|^4}{\mathsf{E}^{\times}(A)}\,. \tag{6.5}$$

By assumption, $|A|^6 \lesssim p^2\mathsf{E}^{\times}(A)$, and so $|A||P||A_1| \lesssim p^2$.

To strengthen the bound on the number of points to $\leq p^2$ (as is required), we partition the set of points in $A \times A_1 \times P$ into $\lesssim 1$ piece $\{\mathcal{P}_i\}$ whose size differs by at most a constant factor, such that each $|\mathcal{P}_i| \leq p^2$. This strengthening adds a $\log|A|$ factor to the estimate of $\mathsf{E}^{+}(A_1)$, which is subsequently hidden by the $\lesssim$ notation.

Having now established the applicability of Rudnev's bound, we turn to the estimate it yields; for this we require a bound on the maximum number of collinear points. This is bounded by $\max(|A|, |P|)$.

Thus,

$$\mathsf{E}^{+}(A_1) \lesssim q^{-2}\left((|A_1||A||P|)^{3/2} + \max(|A|,|P|)|A_1||A||P|\right)\,.$$

We claim that the first term is dominant.
Suppose first that $\max(|A|, |P|) = |P|$. For the first term to dominate, we need to show that $|A||A_1| \geq |P|$; this follows immediately from $|P| \lesssim |A_1|^2$, a consequence of equation (6.5).
Suppose instead that $\max(|A|, |P|) = |A|$. We must show that $|A_1||P| \gtrsim |A|$. Observe that if this is not the case then, since $q \leq |P|$ we return to (6.4) and have that

$$|P||A_1| \geq q|A_1| \gtrsim \frac{\mathsf{E}^{\times}(A)}{|A|} \geq |A|\,,$$

and so $|A_1||P| \gtrsim |A|$, and hence the first term must dominate.

We have now proved that

$$\mathsf{E}^{+}(A_1) \lesssim q^{-2}\,(|A||P||A_1|)^{3/2}\;;$$

to conclude the proof, we use the estimates $q|A_1| \lesssim |P|t$, $\mathsf{E}^{\times}(A) \sim |P|t^2$ and (6.5). Thus

$$\mathsf{E}^{+}(A_1) \lesssim \frac{|A_1|^{7/2}|A|^{3/2}}{|P|^{1/2}t^2} \sim \frac{|A_1|^{7/2}|A|^{3/2}|P|^{1/2}}{\mathsf{E}^{\times}(A)} \lesssim \frac{|A_1|^{11/2}|A|^{3/2}}{\mathsf{E}^{\times}(A)^{3/2}}\,.$$

To obtain the version of the statement in which $\mathsf{E}^{\times}(A)$ is replaced by $\mathsf{E}^{+}(A)$ and $\mathsf{E}^{+}(A_1)$ is replaced with $\mathsf{E}^{\times}(A_1)$, one repeats the same argument almost verbatim, swapping any instance of multiplication by addition, and vice versa. □

The proof in the complex analogue deviates only in the estimate of $\mathsf{E}^{+}(A_1)$, where instead we use the Szemerédi-Trotter theorem.

Proposition 6.9. *Let $A \subseteq \mathbb{C}$ be a finite non-empty set. Then there is $A_1 \subseteq A$ such that $|A_1|^2 \gtrsim \mathsf{E}^{\times}(A)|A|^{-1}$ and*

$$\mathsf{E}^{+}(A_1)\mathsf{E}^{\times}(A) \lesssim |A_1|^{9/2}|A|\,. \tag{6.6}$$

The energies $\mathsf{E}^{\times}$ and E^{+} may be swapped.

Proof. Suppose we have the set $A_1 \subseteq A$ already from the same procedure as in the proof of Proposition 6.8. We know that A_1 satisfies $|A_1|^2|A| \gtrsim \mathsf{E}^{\times}(A)$.

Let $S_\tau = \{x \in A_1 + A_1 : \tau \leq r_{A_1+A_1}(x) < 2\tau\}$ be the set of 'rich sums' in A_1.

Introducing a 'richness threshold' T, to be optimised later, and observing that $\sum_x r_{A_1+A_1}(x) \leq |A_1|^2$, we can write

$$\mathsf{E}^{+}(A_1) = \sum_x r^2_{A_1+A_1}(x) \leq T|A_1|^2 + \sum_{\tau \geq T} (2\tau)^2 |S_\tau|\,,$$

where the second sum is over dyadic τ.

Explicitly, $\sum_\tau \cdot = \sum_{i=1}^{\lceil \log |A| \rceil} \sum_{\tau=2^i} \cdot$.

Since we choose A_1 as in Proposition 6.8, we inherit the associated set P, which is the set of slopes in A/A popular with respect to multiplicative energy. Also, each element of A_1 has at least q representations as either A/P or AP (depending on whether we choose A_1 with respect to the ordinates or abscissae).

We estimate the size of S_τ using the Szemerédi-Trotter theorem; since any element in A_1 can be expressed as a product (or ratio) in AP at least q times, we have

$$\begin{aligned}\tau q|S_\tau| &\leq |\{(x,a_1,\alpha,r) \in S_\tau \times A_1 \times A \times P : x = a_1 + \alpha r\}| \\ &= \mathcal{I}(\text{ points } S_\tau \times A_1, |A||P| \text{ lines})\,.\end{aligned}$$

Applying the Szemerédi-Trotter theorem and rearranging yields either that $\tau < q^{-1}|A_1|$ or

$$|S_\tau| \ll \frac{|A|^2|A_1|^2|P|^2}{\tau^3 q^3} + \frac{|A||P|}{\tau q}\,.$$

Note that the second term dominates when $\tau^2 \geq |A_1||A||P|^2 q^{-2}$.

Hence,

$$\begin{aligned}\mathsf{E}^+(A_1) &\leq T|A_1|^2 + \sum_{\substack{i=1 \\ 2^{i-1}T<\tau\leq 2^i T}}^{\log(|A|T^{-1})} \tau^2|S_\tau| + \mathcal{E} \\ &\leq T|A_1|^2 + \sum_i \frac{|A_1|^2|A|^2|P|^2}{q^3 T 2^i} + \frac{|A_1||P|2^i T}{q} + \mathcal{E} \\ &\lesssim T|A_1|^2 + \frac{|A_1|^2|A|^2|P|^2}{q^3 T} + \frac{|A_1||P||A|}{q} + \mathcal{E}\,.\end{aligned}$$

In the above, we set $\mathcal{E} = \sum \tau^2|S_\tau|$, where the sum is taken over dyadic $\tau < q^{-1}|A_1|$ – that is, the regime in which Szemerédi-Trotter is not applicable. This term is an error term; indeed

$$\mathcal{E} \leq q^{-1}|A_1| \sum \tau|S_\tau| = \frac{|A_1|^3|A_1|}{q|A_1|} \leq \frac{|A_1|^4|A|}{\mathsf{E}^\times(A)}\,.$$

This estimate is stronger that what is required in the statement of the proposition.

We have incurred a logarithmic loss (in $|A|$) by bounding the largest dyadic interval. We also use the trivial inequality $\tau \leq |A_1|$. The parameter T is chosen to be $T = |P||A|q^{-3/2}$.

Hence,

$$\mathsf{E}^+(A_1) \lesssim \frac{|A||P||A_1|^2}{q^{3/2}} + \frac{|P||A_1|^2}{q} + \frac{|A_1|^4|A|}{\mathsf{E}^\times(A)}\,. \tag{6.7}$$

To estimate the first term, we use the estimates $|P|t^2 \sim \mathsf{E}^\times(A)$, and $|P|t \lesssim |A_1|^2$ and obtain

$$\frac{|A||P||A_1|^{7/2}}{(q|A_1)^{3/2}} \lesssim \frac{\sqrt{|P|t}|A||A_1|^{7/2}}{|P|t^2} \lesssim \frac{|A_1|^{9/2}|A|}{\mathsf{E}^{\times}(A)}\,.$$

The estimation of the second term proceeds in a similar fashion:

$$\frac{|P||A_1|^2}{q} = \frac{|P|t|A_1|^3}{q|A_1|t} \lesssim \frac{|A_1|^5}{\mathsf{E}^{\times}(A)}\,.$$

This concludes the proof of the proposition. □

Remark 6.10. *In arbitrary fields, we could use the point-line incidence theorem Theorem 4.5 in place of the Szemerédi-Trotter theorem in the proof of Proposition 6.9. This yields, under suitable conditions on the cardinality of $|A|$ in terms of the characteristic of the field, the estimate:*

$$\mathsf{E}^{+}(A_1)\ \mathsf{E}^{\times}(A)^{2/3} \lesssim |A_1|^{11/3}|A|\,.$$

However using this estimate in the forthcoming proof of Theorem 6.6 yields a quantitatively weaker result.

Bootstrapping the initial decomposition and proofs of Theorems 6.5 and 6.6

With Proposition 6.9 to hand, it is tempting to attempt to directly prove Theorem 6.1. That is, given that we are now able to find a set A_1 such that $\mathsf{E}^{+}(A_1)\mathsf{E}^{\times}(A) \lesssim |A_1|^{9/2}|A|$, can one immediately ascertain that there exists a set A' such that $\min\{\mathsf{E}^{+}(A'), \mathsf{E}^{\times}(A') \lesssim |A|^{3-1/4}\}$?

This is too much to hope for, as Theorem 6.1 asks that A' is a positive proportion of A.

However, if we were to relax this condition, setting A' to be the A_1 of Proposition 6.9, Proposition 6.9 would yield

$$\min\{\mathsf{E}^{+}(A'), \mathsf{E}^{\times}(A)\} \lesssim \big(\mathsf{E}^{+}(A')\mathsf{E}^{+}(A)\big)^{1/2} \lesssim |A|^{11/4}\,.$$

To attain the full strength of Theorems 6.5 and 6.6, we pass the set A_1 through the same bootstrapping algorithm that was used by Balog and Wooley [4]. Roughly, we use Propositions 6.8 and 6.9 (depending on the field of interest) to carve out from A a 'large' subset B with 'small' energy. We iterate this procedure on $A \setminus B$, adding the 'small energy' section to the subset B. Eventually the untouched elements of A will either have sufficiently small

multiplicative energy, or be sufficiently small in cardinality (thus automatically having a small multiplicative energy). The additive energy of the set B, which we constructed at each stage by choosing the 'small energy' subset, cannot grow too large, as the following lemma shows:

Lemma 6.11. *Let $A_1, \dots, A_n$ be finite subsets of an abelian group, for some $n \in \mathbb{N}$. Then*

$$\left(\mathsf{E}^+\left(\bigcup_{i=1}^n A_i\right)\right)^{\frac{1}{4}} \leq \sum_{i=1}^n \left(\mathsf{E}^+(A_i)\right)^{\frac{1}{4}} . \tag{6.8}$$

Proof. This follows from two applications of Cauchy-Schwarz:

$$\begin{aligned}
\left(\mathsf{E}^+\left(\bigcup_{i=1}^n A_i\right)\right)^{\frac{1}{4}} &= \sum_{1\leq i,j,k,l\leq n} \sum_x r_{A_i+A_j}(x) r_{A_k+A_l}(x) \\
&\leq \sum_{1\leq i,j,k,l\leq n} \sqrt{\sum_x r^2_{A_i+A_j}(x) \sum_x r^2_{A_k+A_l}(x)} \\
&= \left(\sum_{i,j} \sqrt{\sum_x r^2_{A_i+A_j}(x)}\right)^2 \\
&\leq \left(\sum_{i,j} \left(\sum_i r^2_{A_i+A_i}(x) \sum_x r^2_{A_j+A_j}(x)\right)^{1/4}\right)^2 \\
&= \left(\sum_i \left(\sum_x r^2_{A_i+A_i}(x)\right)^{\frac{1}{4}}\right)^4 .
\end{aligned}$$

□

Proof of Theorems 6.5 and 6.6

We will first prove the general field version, Theorem 6.6.

Let $M \geq 1$ be a parameter to be chosen later. We construct a decreasing sequence of sets

$$C_1 = A \supseteq C_2 \supseteq \cdots \supseteq C_k$$

and an increasing sequence of sets

$$B_0 = \emptyset \subseteq B_1 \subseteq \cdots \subseteq B_{k-1} \subseteq A$$

such that for any $j = 1, 2, \ldots, k$ the sets C_j and B_{j-1} are disjoint and moreover $A = C_j \sqcup B_{j-1}$.

If at some step j we have that $\mathsf{E}^{\times}(C_j) \leq |A|^3/M$, then we stop our algorithm putting $C = C_j$, $B = B_{j-1}$, and $k = j - 1$. Else, we have $\mathsf{E}^{\times}(C_j) > |A|^3/M$. We apply Proposition 6.8 to the set C_j. In order to apply Proposition 6.8 we must have that $|C_j|^6 \ll p^2\mathsf{E}^{\times}(C_j)$. We have

$$\frac{|C_j|^6}{\mathsf{E}^{\times}(C_j)} < \frac{|A|^6}{|A|^3/M} = |A|^3 M\,, \tag{6.9}$$

so we require that $|A|^3 M \ll p^2$. Hence we find the subset D_j of C_j such that

$$|D_j|^2 \gtrsim \frac{\mathsf{E}^{\times}(C_j)}{|C_j|} > \frac{|A|^3}{M|C_j|} \geq \frac{|A|^2}{M}$$

and

$$\mathsf{E}^{+}(D_j) \lesssim \frac{|D_j|^{11/2}|C_j|^{3/2}}{(\mathsf{E}^{\times}(C_j))^{3/2}} \leq \frac{M^{3/2}|D_j|^{11/2}|C_j|^{3/2}}{|A|^{9/2}} < |D_j|^{11/2}M^{3/2}|A|^{-3}\,. \tag{6.10}$$

Then, we set $C_{j+1} = C_j \setminus D_j$, $B_j = B_{j-1} \sqcup D_j$ and iterate this procedure.

Eventually this process will terminate (as $|C_j|$ decreases) after say k iterations. So $B = B_k = \bigsqcup_{j=1}^{k} D_j$ and $C = C_{k+1}$.

Since $|B| \leq |A|$ it follows that $\sum_{j=1}^{k} |D_j| \leq |A|$.
Then, using Lemma 6.11, we have

$$\begin{aligned}(\mathsf{E}^{+}(B_k))^{1/4} &\leq (M^{3/2}|A|^{-3})^{1/4} \sum_{j=1}^{k} |D_j|^{11/8} \\ &\lesssim (M^{3/2}|A|^{-3})^{1/4} \max_j\{|D_j|\}^{3/8} \sum_{j=1}^{k} |D_j| \\ &\lesssim M^{3/8}|A|^{5/8}\,.\end{aligned}$$

Finally, if we choose $M = |A|^{1/5}$, we obtain the result. Note that with this choice of M, we have $|A|^3 M \ll p^2$, and thus we satisfy the constraint required to apply Proposition 6.8 of (6.9). This completes the proof over $\mathbb{F}$.

The corresponding proof over $\mathbb{R}$ is very similar, using Proposition 6.9 in place of Proposition 6.8. The conclusion of the proof of Theorem 6.5 is exactly the algorithmic approach as before, but with the exponents suitably amended.

Proof of Corollary 6.7

To prove the first statement, equation (6.1), of Corollary 6.7, we repeat the proof of Theorem 6.6 verbatim to obtain disjoint sets B_k and C_k such that $\mathsf{E}^+(B_k) \lesssim M^{3/2}|A|^{5/2}$ and $\mathsf{E}^\times(C_k) \leq |A|^3 M^{-1}$. We then estimate

$$\mathsf{E}^+(B_k)\mathsf{E}^\times(C)^{3/2} \lesssim M^{3/2}|A|^{5/2}|A|^{9/2}M^{-3/2} = |A|^7\,,$$

as required.

The proof of equation (6.2) is similar: using the same B_k and C_k as above, we find that

$$\mathsf{E}^+(B_k)\mathsf{E}^\times(C_k) \lesssim M^3|A|^5 + \frac{|A|^6}{M^2}\,,$$

where in the first term of the estimate, we have assumed that $\mathsf{E}^+(B_k)$ is the bigger sum, and the second term is assuming the converse. We choose $M = |A|^{1/5}$. The improvement in the real case is a consequence of using the stronger Proposition 6.9 in the algorithmic proof of Theorem 6.5.

7

Applications of energy decomposition

7.1 Road-map of this chapter

In this chapter we will consider two applications of the energy formulation of the sum-product conjecture.

The first application is to expansion. We will motivate expansion and prove a four-variable expansion result in finite fields, using the energy decomposition result as a tool.

The second application of the energy reformulation of the sum-product problem to the sum-product problem itself. We will state and prove a new bound. The literature associated with this bound uses the notion of *Szemerédi-Trotter sets*, and associated notation; we provide a full proof, which we hope will be a useful resource.

7.2 Application: Expansion

Expander functions

A consequence of the decomposition of a set into additive and multiplicative energy is that of expansion: imagine that you have a black box machine which takes a number of inputs, and combines them using only the operations multiplication and addition. For example, consider the real-valued two-variable function $f(x, y) = x(x + y)$. Then, using the philosophy that multiplication

and addition do not coexist, no matter how we restrict the input variables that we feed this black box, the output must be 'large'. For our illustrative f, this translates into asking whether the following is true:

$$\text{For any } A \subseteq \mathbb{R},\ |\{f(a,b)\colon a,b \in A\}| \gg |A|^{1+\epsilon} \text{ for } 0 < \epsilon < 1.$$

A function which obeys this expansion property is called an *expander*. More precisely:

Definition 7.1. *Let $d \geq 2$ be a positive integer, and $\mathbb{F}$ be a field. A function $f : \mathbb{F}^d \to \mathbb{F}$ is called a d-dimensional expander if there exists $0 < \epsilon < d-1$ such that for any finite (non-empty) set $A \subset \mathbb{F}$ we have*

$$|\{f(a_1, \ldots, a_d\} : a_i \in A\}| \gg |A|^{1+\epsilon}.$$

The goal of expansion is to find a function for which we can prove an explicit value of ϵ as large as possible.

The applications of expanders to pseudorandomness are briefly discussed towards the end of Section 5.2. From an additive combinatorial viewpoint, the aim of expansion is to achieve the highest rate of expansion (i.e. to maximise ϵ/d). Motivated by the pseudorandomness applications, small values of d are of particular interest.

In this chapter we will limit ourselves to the consideration of expansion over the prime reside field $\mathbb{F}_p$. In this context, the question of expansion can be rephrased as follows: given a function $f : \mathbb{F}_p^d \to \mathbb{F}_p$, how large (in terms of cardinality) must a set A be until

$$|\{f(a_1, \ldots, a_d) : a_i \in A\}| \gg p\,?$$

For some choices of f, there exist obstructions preventing the $\gg$ being an equality – see, for example Theorem 8.8. In this rephrasing of the question, the goal is now to take $|A|$ to be as *s*mall as possible until a positive proportion of the field is covered by $f(A)$.

In this situation, for $d = 3, 4$, there are a number of instances of functions f achieving the 'threshold bound': if $|A| \gg p^{2/3}$ then $|f(A)| \gg p$. When $d = 3$, Pham, Vinh and de Zeeuw [83, Theorem 1.1] proved this threshold bound for a large class of functions f lacking a particular structure. To illustrate this when $d = 4$, we give the example $f(x_1, x_2, x_3, x_4) = x_1x_2 + x_3x_4$ [89, 37, 45], and refer to [68, 82] for numerous further explicit examples when $d = 3, 4$.

Techniques using character sums or often just linear algebra methods often work well for relatively large sets with respect to p; in the literature, often the question is able to be manipulated into an incidence bound optimised for the realm of Vinh's [116] bound. However, these techniques usually fail to work for smaller A, and it is this gap in the literature that we address.

A new bound on a four-variable expander

In four variables, there are few examples of 'threshold-breaking' bounds. One such is a result of Vinh [115], and the second is by Petridis [82]. They both proved results of the type: if $|A| \geq p^{5/8}$, then $|f(A)| \gg p$. Vinh used the function $f(x_1, x_2, x_3, x_4) = x_1x_2 + (x_3 - x_4)^2$ and graph-theoretic methods; Petridis studied the functions $f(x_1, x_2, x_3, x_4) = (x_1 \pm x_2)(x_3 \pm x_4)$ using combinatorial methods. To achieve his bound, Petridis uses a second moment estimate on additive energy of a set A and its dilate, in essence, bounding the variance of this energy from its expected mean.

In this section, we pass beyond the exponent of $5/8$, using the method of Petridis and an improvement that is a consequence of the energy decomposition described in Chapter 6.

Theorem 7.2. *Let $A \subseteq \mathbb{F}_p$. The number of solutions to the equation*

$$\frac{ab - c}{a - d} = \frac{a'b' - c'}{a' - d'} \tag{7.1}$$

with $a, b, c, d, a', b', c', d' \in A$ is $\frac{|A|^8}{p} + O(p^{2/3}|A|^{79/15})$.
Hence if $|A| = \Omega(p^{25/42})$ then

$$\left|\left\{\frac{ab - c}{a - d} : a, b, c, d \in A\right\}\right| = \Omega(p)\,.$$

.

We conclude this section with a remark on the current best expansion result by Murphy et al. [70]. Using the three-variable function $f(a, b, c) := \frac{b-a}{c-a}$, first studied over $\mathbb{R}$ by Jones [52], they show that if $|A| \geq p^{3/5}$, then $f(A) \gg p$.

Proof of Theorem 7.2

The proof of Theorem 7.2 begins by restricting the count of solutions of the type (7.1)) to the number of solutions to the equation

$$\frac{a_1b_1 - c_2}{a_1 - d_2} = \frac{a_1'b_1' - c_2'}{a_1' - d_2'} \quad a_1, b_1, a_1', b_1' \in A_1, c_2, d_2, c_2', d_2 \in A_2$$

where $A_1, A_2, \subseteq A$ are disjoint of size $|A_1|, |A_2| \geq |A|/3$.

The sets A_1 and A_2 are chosen according to the Balog-Wooley analogue in $\mathbb{F}_p$ – Theorem 6.6. What is important is that a suitably weighted combination of the product of $\mathsf{E}^{\times}(A_1)$ and $\mathsf{E}^{+}(A_2)$ will be small.

We will use the more convenient notation

$$\left|\left\{\frac{a_1b_1 - c_2}{a_1 - d_2} = \frac{a_1'b_1' - c_s'}{a_1 - d_s'}\right\}\right|$$

to refer to the number of solutions of the equation from interest whenever it is clear in which sets the variables lie.

We begin the proof with a standard application of the Cauchy-Schwarz inequality:

$$\begin{aligned}|A_1|^4|A_2|^4 &= \left(\sum_{x\in\mathbb{F}_p}\left|\left\{\frac{a_1b_1 - c_2}{a_1 - d_2} = x\right\}\right|\right)^2 \\ &\leq \left|\left\{\frac{a_1b_1 - c_2}{a_1 - d_2} \colon a_1, b_1 \in A_1, c_2, d_2 \in A_2\right\}\right| \left|\left\{\frac{a_1b_1 - c_2}{a_1 - d_2} = \frac{a_1'b_1' - c_2'}{a_1' - d_2'}\right\}\right| .\end{aligned}$$

The second term is the L_2 version of the object of interest, measuring how many pairs of quadruples coincide; it is the energy expression for this equation.

We have

$$\begin{aligned}\left|\left\{\frac{a_1b_1 - c_2}{a_1 - d_2} = \frac{a_1'b_1' - c_2'}{a_1' - d_2'}\right\}\right| &\leq |A_1|^3|A_2|^3 + \sum_{x\neq 0}\left|\left\{\frac{a_1b_1 - c_2}{a_1 - d_2} = x = \frac{a_1'b_1' - c_2'}{a_1' - d_2'}\right\}\right| \\ &\leq |A_1|^3|A_2|^3 + \sum_{x\neq 0}|\{a_1(b_1 - x) = c_2 - d_2x\}|^2 \\ &\leq |A_1|^3|A_2|^3 + \sum_{x\neq 0}\mathsf{E}^{\times}(A_1, A_1 + x)\mathsf{E}^{+}(A_2, xA_2)\,.\end{aligned}$$

We rearrange this to a somewhat curious-looking expression, the motivation

for which will become clear:

$$\begin{aligned}
&\sum_{x\neq 0} \mathsf{E}^{\times}(A_1, A_1+x)\mathsf{E}^{+}(A_2, xA_2) \\
&\quad = \sum_{x\neq 0}\left(\mathsf{E}^{\times}(A_1, A_1+x) - \frac{|A_1|^4}{p}\right)\left(\mathsf{E}^{+}(A_2, xA_2) - \frac{|A_2|^4}{p}\right) \\
&\quad + \frac{|A_2|^4}{p}\sum_{x\neq 0}\mathsf{E}^{\times}(A_1, A_1+x) + \frac{|A_1|^4}{p}\sum_{x\neq 0}\mathsf{E}^{+}(A_2, xA_2) - \frac{p-1}{p^2}|A_1|^4|A_2|^4 \\
&\quad \leq \left(\sum_{x\neq 0}\left(\mathsf{E}^{\times}(A_1, A_1+x) - \frac{|A_1|^4}{p}\right)^{1+s}\right)^{\frac{1}{1+s}}\left(\sum_{x\neq 0}\left(\mathsf{E}^{+}(A_2, xA_2) - \frac{|A_2|^4}{p}\right)^{1+t}\right)^{\frac{1}{1+t}} \\
&\quad + \frac{|A_2|^4}{p}\sum_{x\neq 0}\mathsf{E}^{\times}(A_1, A_1+x) + \frac{|A_1|^4}{p}\sum_{x\neq 0}\mathsf{E}^{+}(A_2, xA_2) - \frac{p-1}{p^2}|A_1|^4|A_2|^4,
\end{aligned}$$

where $\frac{1}{1+s} + \frac{1}{1+t} = 1$ and we have used Hölder's inequality. We choose to present $s, t > 0$ as parameters to be optimised in the course of the proof. It will turn out that we choose $s = 3/2$ and $t = 2/3$.

We justify the omission of absolute value signs in the above application of Hölder via an application of Cauchy-Schwarz's inequality: for any set $B \subseteq \mathbb{F}_p$ and $\xi \in \mathbb{F}_p^*$, we have

$$|B+\xi B|\mathsf{E}^{+}(B, \xi B), |B(\xi+B)|\mathsf{E}^{\times}(B, \xi+B) \geq |B|^4; \tag{7.2}$$

since $|B+\xi B|, |B(\xi+B)| \leq p$, (or indeed, $|\mathbb{F}|$ whenever $\mathbb{F}$ is finite) we find that, e.g., $\mathsf{E}^{\times}(A, x+A) - \frac{|A|^4}{p} > 0$.

The reasoning for the above expression is to take advantage of an argument of Petridis [82], modifying it to suit our needs. Although Petridis considered $\mathsf{E}^{+}(A, xA)$ in his work, his methods extend to the analogous $\mathsf{E}^{\times}(A, x+A)$.

Our main amendment to Petridis' argument is a change to Proposition 8 in [82] which we present in the following form.

Proposition 7.3. *Let $A \subseteq \mathbb{F}_p$ and let $s \in (0, 3)$. Then*

$$\sum_{x\neq 0}\left(\mathsf{E}^{+}(A, xA) - \frac{|A|^4}{p}\right)^{1+s} = O\left(p^{1-\frac{s}{3}}\ \mathsf{E}^{+}(A)^{\frac{2s}{3}}\ |A|^{2+\frac{4s}{3}}\right),$$

and

$$\sum_{x\neq 0}\left(\mathsf{E}^{\times}(A, x+A) - \frac{|A|^4}{p}\right)^{1+s} = O\left(p^{1-\frac{s}{3}}\ \mathsf{E}^{\times}(A)^{\frac{2s}{3}}\ |A|^{2+\frac{4s}{3}}\right),$$

The proof of this relies on the following explicit version of Bourgain's Theorem C from [14] as communicated in [82] (Theorem 2 and 5) in additive and multiplicative form by Rudnev and Murphy respectively.

Lemma 7.4. *Let $A \subseteq \mathbb{F}_p$ and $X \subseteq \mathbb{F}_p^*$. Suppose that $|A|^2|X| = O(p^2)$. Then*

$$\sum_{x\in X} \mathsf{E}^+(A, xA) = O\left(\mathsf{E}^+(A)^{1/2}\left(|A|^{3/2}|X|^{3/4} + |A||X|\right)\right)$$

and

$$\sum_{x\in X} \mathsf{E}^\times(A, x+A) = O\left(\mathsf{E}^\times(A)^{1/2}\left(|A|^{3/2}|X|^{3/4} + |A||X|\right)\right).$$

Sketch. The main idea of this proof is to use Rudnev's point-plane incidence bound, Theorem 2.9. The quantity $\sum_{x\in X} \mathsf{E}^+(A, xA)$ is the count of solutions to

$$a - c = x(d - b) \quad a, b, c, d \in A; x \in X\,;$$

by Cauchy-Schwarz, this can be bounded by $\sqrt{\mathsf{E}^+(A)}$ multiplied by the (square root of the) number of solutions to

$$x(d - b) = x'(d' - b') \quad b, d, b', d' \in A; x, x' \in X\,,$$

which can be realised as an incidence bound between $|A|^2|X|$ planes and the same number of points.

The multiplicative version is analogous, but we instead count solutions to

$$\frac{a}{c} = \frac{x+b}{x+c}\,.$$

This is bounded by $\sqrt{\mathsf{E}^\times(A)}$ multiplied by the (square root of the) number of solutions to

$$\frac{x+b}{x+c} = \frac{x'+b'}{x'+c'} \quad b, b', c, c' \in A; x, x' \in X\,.$$

We interpret this as an incidence bound between $|X||A|^2$ planes of the form $\{(u, v, w) \in \mathbb{F}^3 \colon xu + bv - cw = 0\}$ with $x \in X$ and $b, c \in A$, and at most $|A|^2|X|$ points of the form $(u, v, w) = (c' - b', x' + c', x' + b') \in \mathbb{F}^3$ with $b', c' \in A$ and $x' \in X$. □

We note that variants of Lemma 7.4 can easily be obtained by using a point-line incidence bound in place of a points-planes incidence bound. However, for our purposes, the stated version is the most advantageous.

We will also need the following version of Lemma 4 from [82], which follows from Lemma 7.4 by setting $X = \{x \in \mathbb{F}_p^* : E^+(A, xA) > \mathsf{E}^+(A)K^{-1}\}$.

Lemma 7.5. *Suppose that $A \subseteq \mathbb{F}_p$ such that $|A| \geq p^{1/2}$ and $K|A|^4 \ll p\,\mathsf{E}^+(A)$.*

Then there are at most $O\left(K^4|A|^6\mathsf{E}^+(A)^{-2}\right)$ elements $x \in \mathbb{F}_p^*$ satisfying $\mathsf{E}^+(A, xA) > \frac{\mathsf{E}^+(A)}{K}$.

The multiplicative version also holds, swapping $\mathsf{E}^+(A)$ for $\mathsf{E}^\times(A)$. Note that K is well-defined from (7.2).

We quote an auxiliary lemma of Bourgain [17] which appears (and is proved) explicitly in [82] as Lemma 3.

Lemma 7.6. *For every set $X \subseteq \mathbb{F}_p^*$ we have the following inequalities*

$$\sum_{x\in X}\left(\mathsf{E}^+(A, xA) - \frac{|A|^4}{p}\right), \sum_{x\in X}\left(\mathsf{E}^\times(A, x+A) - \frac{|A|^4}{p}\right) \leq p|A|^2\,.$$

We now prove Proposition 7.3; the proofs for both statements are almost identical so we prove only the additive version.

Proof of Proposition 7.3. We will bound

$$\sum_{x\neq 0}\left(\mathsf{E}^+(A, xA) - \frac{|A|^4}{p}\right)^{1+s}\,.$$

Let $K > 0$ be a parameter measuring $\mathsf{E}^+(A, xA)$, to be determined later.

The case when $\mathsf{E}^+(A, xA)$ is small is dealt with via Lemma 7.6:

$$\begin{aligned}\sum_{\substack{x\neq 0\\ \mathsf{E}^+(A,xA)\leq\frac{\mathsf{E}^+(A)}{K}}}\left(\mathsf{E}^+(A, xA) - \frac{|A|^4}{p}\right)^{1+s} &\leq \sum_{\substack{x\neq 0\\ \mathsf{E}^+(A,xA)\leq\frac{\mathsf{E}^+(A)}{K}}}(\mathsf{E}^+(A, xA))^s\left(\mathsf{E}^+(A, xA) - \frac{|A|^4}{p}\right)\\ &\leq \sum_{\substack{x\neq 0\\ \mathsf{E}^+(A,xA)\leq\frac{\mathsf{E}^+(A)}{K}}}\left(\mathsf{E}^+(A, xA) - \frac{|A|^4}{p}\right)\ \left(\frac{\mathsf{E}^+(A)}{K}\right)^s\\ &\leq p|A|^2\left(\frac{\mathsf{E}^+(A)}{K}\right)^s\,.\end{aligned}$$

For 'large' $\mathsf{E}^+(A, xA)$ we use a dyadic argument.
Let

$$X_i := \{x \neq 0 : \mathsf{E}^+(A)2^{-i} \leq \mathsf{E}^+(A, xA) < \mathsf{E}^+(A)2^{1-i}\}$$

for $i = 1, \ldots, k$, where $2^{k-1} < K \leq 2^k$.

By Lemma 7.5 we know that $|X_i| = O\left(2^{4i}\frac{|A|^6}{E_+(A)^2}\right)$. So

$$\begin{aligned}\sum_{\substack{\mathsf{E}^+(A,xA)>\\ >\mathsf{E}^+(A)/K}} \left(\mathsf{E}^+(A,xA) - \frac{|A|^4}{p}\right)^{1+s} &\leq \sum_{\substack{\mathsf{E}^+(A,xA)>\\ \frac{\mathsf{E}^+(A)}{K}}} \mathsf{E}^+(A,xA)^{1+s}\\ &= \sum_{i=1}^{k}\sum_{x\in X_i}\mathsf{E}^+(A,xA)^{1+s}\\ &\ll \sum_{i=1}^{k}|X_i|\mathsf{E}^+(A)^{1+s}2^{-(1+s)i}\\ &= O(K^{3-s}\mathsf{E}^+(A)^{s-1}|A|^6)\end{aligned}$$

Finally, we optimise by combining the 'small energy' and the 'large energy case', choosing $K^3 = p\ \mathsf{E}^+(A)|A|^{-4}$. □

We now return to the estimation of $\sum_{x\neq 0}\mathsf{E}^\times(A_1, x+A_1)\mathsf{E}^+(A_2,xA_2)$, using Proposition 7.3 in the previously-found rearrangement, recalling that $\frac{1}{1+s}+\frac{1}{1+t}=1$ and $|A_1|,|A_2|\leq|A|$:

$$\begin{aligned}&\sum_{x\neq 0}\mathsf{E}^\times(A_1,xA_1)\mathsf{E}^+(A_2,xA_2)\\ &\quad\ll \left(p^{1-\frac{s}{3}}\ \mathsf{E}^\times(A_1)^{\frac{2s}{3}}\ |A_1|^{2+\frac{4s}{3}}\right)^{\frac{1}{1+s}}\left(p^{1-\frac{t}{3}}\ \mathsf{E}^+(A_2)^{\frac{2t}{3}}\ |A_2|^{2+\frac{4t}{3}}\right)^{\frac{1}{1+t}} + \frac{|A_1|^4|A_2|^4}{p}\\ &\quad\ll p^{2/3}|A|^{10/3}\left(\mathsf{E}^\times(A_1)^{\frac{s}{1+s}}\mathsf{E}^+(A_2)^{\frac{t}{1+t}}\right)^{2/3} + \frac{|A|^8}{p}\\ &\quad\ll p^{2/3}|A|^{10/3}\left(\mathsf{E}^+(A_2)\mathsf{E}^\times(A_1)^s\right)^{\frac{2}{3(1+s)}} + \frac{|A|^8}{p}\end{aligned}$$

We now apply Corollary 6.7 to the energy term, choosing $s=3/2$; hence we find that $\mathsf{E}^+(A_2)\mathsf{E}^\times(A_1)^{3/2}\ll|A|^7$, and so

$$\sum_{x\neq 0}\mathsf{E}^\times(A_1,x+A_1)\mathsf{E}^+(A_2,xA_2)\lesssim\frac{|A|^8}{p}+p^{2/3}|A|^{26/5}\,.$$

Hence, using this and Lemma 7.4, we bound

$$\sum_{x\neq 0}\mathsf{E}^{\times}(A_1,A_1+x)\mathsf{E}^{+}(A_2,xA_2) \ll p^{2/3}|A|^{10/3}\left(\mathsf{E}^{+}(A_2)\mathsf{E}^{\times}(A_1)^{s}\right)^{\frac{2}{3(1+s)}} + \frac{|A|^8}{p}$$
$$+ \frac{|A_2|^4}{p}\mathsf{E}^{\times}(A_1)(|A_1|^{3/2}p^{3/4}+p|A_1|)$$
$$+ \frac{|A_1|^4}{p}\mathsf{E}^{+}(A_2)(|A_2|^{3/2}p^{3/4}+p|A_2|)$$
$$+ \frac{|A_1|^4|A_2|^4}{p}$$

When $|A| \gtrsim p^{25/42}$, the term $\frac{|A|^8}{p}$ dominates, and so, by Cauchy-Schwarz, the set $\{\frac{ab-c}{a-d} : a,b,c,d \in A\}$ occupies a positive proportion of $\mathbb{F}_p$. This concludes the proof of Theorem 7.2.

7.3 The Sum Product Phenomenon

Recall that the sum-product conjecture is that for any non-empty finite set $A \subseteq \mathbb{R}$, we have that $\max(|AA|,|A+A|) \gg |A|^{2-\epsilon}$ for all $\epsilon > 0$. We have previously (Section 3.3) seen how Solymosi [104] connected the multiplicative energy of a set with the cardinality of its sumset, to prove the bound

$$\mathsf{E}^{\times}(A) \lesssim |A+A|^2 . \tag{7.3}$$

Solymosi's argument is both beautiful and elegant but is also wasteful in ways that Konyagin and Shkredov [60, 58] manage to exploit, culminating in the following theorem.

Theorem 7.7. *Let $A \subseteq \mathbb{R}$ be a finite set. Then*

$$\max(|A+A|,|AA|) \gtrsim |A|^{\frac{4}{3}+\frac{5}{9813}} .$$

They also prove stronger results for when $|AA|$ is small, but we will not discuss this here.

In Solymosi's argument, elements in $(A+A)\times(A+A)$ only arise if they can be created as a sum of two consecutive slopes. Whilst this has the advantage that none of these elements coincide, the number of these elements is also a rudimentary lower bound for $|A+A|$.

Shkredov and Konyagin begin with the same picture as Solymosi (Figure 3.3) but instead of creating new vector sums from only *consecutive* slopes, they

group the slopes into bunches, and consider all vector sums created from pairs of points lying on slopes in a bunch. There is no interaction between different bunches, but now it is possible for an element in $(A+A)\times(A+A)$ to have originated in many different ways. Bounding this number of 'collisions' is the key difficulty in Konyagin and Shkredov's argument.

Konyagin and Shkredov used a quantity $d_*(A)$ in their work:

$$d_*(A) := \min_{t>0} \min_{\emptyset\neq Q,R\subseteq\mathbb{R}\setminus\{0\}} \frac{|Q|^2|R|^2}{|A|t^3}$$

where any element $a\in A$ has at least t representations as $a=qr$ for $(q,r)\in Q\times R$ and $\max(|Q|,|R|)\geq|A|$. The idea behind the quantity $d_*(A)$ is that it is a measure of how efficiently the Szemerédi-Trotter theorem can be used in conjunction with the set A. However, this quantity is not overly intuitive, and so we sketch a proof of Konyagin and Shkredov's technique avoiding this quantity.

We also improve Theorem 7.7 using the energy variant of the sum-product phenomenon. This result appears in the paper with Rudnev and Shkredov [92].

Theorem 7.8. *Let $A\subseteq\mathbb{R}$ be a finite set. Then*

$$\max(|A+A|,|AA|)\gtrsim|A|^{\frac{4}{3}+\frac{1}{1509}}.$$

Subsequent Improvements

The current best result on the sum-product phenomenon is by Shakan [97].

Theorem 7.9 (Shakan [97])**.** *Let $A\subseteq\mathbb{R}$ be a finite set. Then*

$$\max(|A+A|,|AA|)\gtrsim|A|^{\frac{4}{3}+\frac{5}{5277}}.$$

Shakan's improvement is the consequence of a decomposition result in the style of Theorem 6.5, but via a close cousin of the quantity $d_*(A)$ and its analogously additively-defined $d_+(A)$. Instead of using $\mathsf{E}+(A)=\mathsf{E}_2(A)$, Shakan turns to $\mathsf{E}_4(A)$ and adapts the proof of Shkredov and Konyagin to better suit this quantity.

Proof of Theorem 7.8

We recall the notation $r_{XY}(a) := |\{(x,y) \in X \times Y : a = xy\}|$. In order to prove Theorem 7.8, we require the following two auxiliary results.

The first auxiliary result replaces the the notation $d_*(A)$ and will be proved in Section 7.3.

Theorem 7.10. *Let $A \subseteq \mathbb{R}\setminus\{0\}$. Suppose there exist non-empty sets $Q, R \subseteq \mathbb{R}$ so that $r_{QR}(a) \geq t$ for all $a \in A$ where $|Q| \geq \max(|A|,|R|)$.*

Then

$$|A+A| \gtrsim |A|^{79/37}\left(\frac{|Q|^2|R|^2}{t^3}\right)^{-21/37}.$$

The second auxiliary result that we require is an energy decomposition result.

Theorem 7.11. *Let $A \subseteq \mathbb{C}$. Then there exists a subset $B \subseteq A$ with $|B| \geq |A|/3$ such that*

$$|B/B| \gtrsim \frac{\mathsf{E}^+(A)^3}{|B|^4|A|^3} \tag{7.4}$$

The skeleton of the proof of Theorem 7.11 is identical to that of say Theorem 6.5 or Corollary 6.7, and so we do not repeat this argument for a third time, deferring instead to the original publication [92, Theorem 2.13] for an explicit exposition of this theorem.

Proof of Theorem 7.8

Let $|A+A| = K|A|$ and $|AA| = M|A|$. Our goal is to show

$$\max(K,M) \gtrsim |A|^{\frac{1}{3}+\frac{1}{1509}}.$$

Without loss of generality[1], suppose that $A \subseteq \mathbb{R}_{>0}$ and consider the set of points $A \times A$.

Step 1: Regularisation

The ratio set A/A is the set of slopes supporting $A \times A$. We will choose a subset of slopes supporting a positive proportion of the multiplicative energy of A. Recall that the multiplicative energy is the sum, over $\lambda \in A/A$, of the

[1]This assumption is further justified in Section 3.3.

number of pairs of points of $A \times A$ lying on the line through the origin with slope λ.

By the popularity principle, we first restrict to the set of points supported on slopes $\lambda \in A/A$ where the line through the origin with slope λ contains at least $\mathsf{E}^{\times}(A)/(2|A|^2)$ points of $A \times A$. Let Λ denote these 'popular' slopes. To see that Λ supports a positive proportion of $\mathsf{E}^{\times}(A)$ note that:

$$\begin{aligned}\mathsf{E}^{\times}(A) &= \sum_{\lambda\in\Lambda} r^2_{A/A}(\lambda) + \sum_{\lambda\notin\Lambda} r^2_{A/A}(\lambda) \\ &\leq \sum_{\lambda\in\Lambda} r^2_{A/A}(\lambda) + |A|^2 \frac{\mathsf{E}^{\times}(A)}{2|A|^2}\,.\end{aligned}$$

Then, by the dyadic pigeonhole principle applied to Λ, there exists a number $|A| \geq \tau \geq \mathsf{E}^{\times}(A)/(2|A|^2)$ and a set $S_\tau \subseteq \Lambda$ such that $|S_\tau|\tau^2 \gg \frac{\mathsf{E}^{\times}(A)}{\log|A|}$. Moreover every $\lambda \in S_\tau$ corresponds to a line through the origin of slope λ supporting between τ and 2τ points of $A \times A$.

Note that if $|S_\tau| \ll 1$, then $|A|^2 \geq \tau^2 = \tau^2|S_\tau||S_\tau|^{-1} \gg \tau^2|S_\tau| \gtrsim \mathsf{E}^{\times}(A) \geq |A|^4|AA|^{-1}$. Whence $M \gtrsim |A|$ and so we are done. In the subsequent arguments we therefore assume that $|S_\tau| \gg 1$.

Step 2: Solymosi's argument and bunches

As in Section 3.3, we order the elements of S_τ consecutively to create elements of $(A+A)\times(A+A)$ from vector sums of points on consecutive slopes. Instead of considering immediately neighbouring slopes, we partition S_τ into bunches of size $2 \leq B \leq |S_\tau|$.

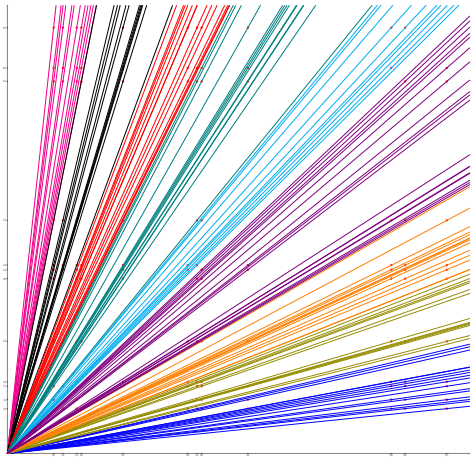


Figure 7.1: We partition lines with slope in S_τ into bunches of B consecutive slopes.

Consider the j-th bunch of lines U_j and the set of points $\mathcal{P}_j \subseteq A \times A$ lying on lines in this bunch. We will consider all vector sums of distinct pairs of points in P_j. Let $A_{\lambda_i} := A \cap \lambda_i A$ be the set of values of the coordinates for the points in $A \times A$ lying on the line with slope λ_i. We abuse notation and associate the line with the slope.

We now use the inclusion-exclusion principle to count the number of elements of $(A+A)\times(A+A)$ which originate from $P_j + P_j$. Note that the geometry of the argument ensures that vector sums originating from the j-th bunch are distinct from the vector sums originating from the k-th bunch for $j \neq k$.

The number of points of $(A+A)\times(A+A)$ that originate from the j-th bunch is by inclusion-exclusion is at least:

$$\rho_j := \tau^2\binom{B}{2} - \sum_{\substack{\lambda_1,\ldots,\lambda_4 \in U_j \\ \lambda_1 \neq \lambda 2, \lambda_3 \neq \lambda 4, \{\lambda_1,\lambda_2\} \neq \{\lambda_3,\lambda_4\}}} \sum_{a_1 \in A_{\lambda_1}} \sum_{a_2 \in A_{\lambda_2}} \sum_{a_3 \in A_{\lambda_3}} \mathbb{1}_{(\lambda_1-\lambda_4)a_1 + (\lambda_2-\lambda_4)a_2 - (\lambda_3-\lambda_4)a_3 = 0}$$

That is, we consider $\tau^2\binom{B}{2}$ elements in $(A+A)\times(A+A)$ but then must exclude the coincidences that we have over-counted. It is readily verified that this is precisely the second term in the expression for ρ_j, and we refer to [60] for this explicit calculation.

We set $\alpha = \frac{\lambda_4-\lambda_2}{\lambda_4-\lambda_1}$ and $\beta = \frac{\lambda_3-\lambda_4}{\lambda_1-\lambda_4}$, and hence

$$\rho_j \geq \tau^2\binom{B}{2} - \sum_{\lambda_1,\lambda_2,\lambda_3,\lambda_4} \sum_{a\in A_{\lambda_1}} r_{\alpha A_{\lambda_2}+\beta A_{\lambda_3}}(a)$$
$$\geq \tau^2\binom{B}{2} - B^4(2\tau \max_{\lambda\in U_j} \mathsf{E}^+(A_\lambda))^{1/2}.$$

In the final inequality, we use the Cauchy-Schwarz inequality. We remark that the subsequent improvement to the sum-product problem by Shakan [97] comes from an application of Hölder's inequality in place of the Cauchy-Schwarz inequality here, and benefiting from its better bounds on the third energy E_3.

We now equipartition the set S_τ into S'_τ and S''_τ so that $\mathsf{E}^+(A_{\lambda'}) \geq \max_{\lambda''\in S''_\tau} \mathsf{E}^+(A_{\lambda''})$ for each $\lambda' \in S'_\tau$.

Let E denote the maximal value of $\mathsf{E}^+(A_{\lambda''})$ among all slopes $\lambda'' \in S''_\tau$.

We have $|A+A|^2 \geq \sum_j \rho_j$ and so, restricting to slopes $\lambda \in S''_\tau$, we have

$$|A+A|^2 \geq \frac{|S''_\tau|}{2B}\left(\tau^2\binom{B}{2} - B^4(2\tau)^{1/2}\mathsf{E}^{1/2}\right)$$
$$\geq \frac{|S_\tau|}{4B}\left(\frac{\tau^2B^2}{4} - B^4(2\tau)^{1/2}\mathsf{E}^{1/2}\right).$$

We choose $B^2 = \frac{1}{2}\lfloor\frac{\tau^{3/2}}{\sqrt{32\mathsf{E}}}\rfloor$ and assume that $B > 1$. If not, then $\mathsf{E} \gg \tau^3$ which is stronger than the required forthcoming (7.5).

Hence

$$\mathsf{E}^{1/4}|A+A|^2 \gg \tau^2|S_\tau|\tau^{3/4}.$$

Using the (multiplicative version of the) Cauchy-Schwarz inequality Lemma 3.11, we obtain

$$\mathsf{E} \gg \frac{\tau^{11}|S_\tau|^4}{|A|^8K^8} \gtrsim \frac{\tau^3|A|^4}{K^8M^4}. \tag{7.5}$$

We now proceed by taking slopes $\lambda \in S'_\tau$, noting that $\mathsf{E}^+(A_\lambda) \geq \mathsf{E}$ for each $\lambda \in S'_\tau$.

Step 3: Relating E and A_λ/A_λ

We now apply Theorem 7.11, replacing A in the statement of Theorem 7.11 with the set A_λ.

Hence there is a subset $\tilde{A}_\lambda \subseteq A_\lambda$ such that $|\tilde{A}_\lambda| \gg |A_\lambda|$ and

$$|\tilde{A}_\lambda/\tilde{A}_\lambda| \gtrsim \frac{\mathsf{E}^+(A_\lambda \setminus \tilde{A}_\lambda)^3}{|\tilde{A}_\lambda|^4|A_\lambda \setminus \tilde{A}_\lambda|^3} \gtrsim \frac{\mathsf{E}^3}{|\tilde{A}_\lambda|^4|A_\lambda|^3} \geq \frac{\mathsf{E}^3}{|A_\lambda|^7} \gg \frac{\tau^{26}|S_\tau|^{12}}{|A|^{24}K^{24}} \tag{7.6}$$

Step 4: Finding a suitable subset of A

Let $\Pi = AA$. Note that from Step 3, every $\lambda \in S'_\tau$ can be represented as a product of two elements from Π in at least $t := |\tilde{A}_\lambda / \tilde{A}_\lambda|$ ways.

By the (ordinary) pigeonhole principle, there exists $a \in A$ such that $A' := A \cap aS'_\tau$ such that $|A'| \gg \tau |S'_\tau||A|^{-1}$ and every element in A' can be written as a product of $a\Pi$ and Π at least t ways.

Hence, by applying Theorem 7.10 with $Q = R = \Pi$, we obtain a lower bound on $|A' + A'|$.

Step 5: Putting it all together

The remainder of the proof of Theorem 7.8 is now a calculation, involving the Cauchy-Schwarz inequality, and bounds on τ and $\tau|S'_\tau|$.

$$\begin{aligned} K^{37}|A|^{37} = |A+A|^{37} \geq |A'+A'|^{37} &\gtrsim \frac{|A'|^{79}t^{63}}{|\Pi|^{84}} \\ &\gtrsim \frac{\tau^{79}|S_\tau|^{79}}{|A|^{79}} \frac{1}{M^{84}|A|^{84}} \frac{\tau^{1638}|S_\tau|^{756}}{|A|^{1512}K^{1512}} \\ &\gtrsim \frac{(\mathsf{E}^\times(A))^{835}\tau^{47}}{M^{84}|A|^{1675}K^{1512}} \gg \frac{|A|^{877}}{M^{966}K^{1512}} \end{aligned}$$

Proving Theorem 7.10

It remains to prove Theorem 7.10. The remainder of this chapter is somewhat technical.

We begin with an auxiliary result summarising properties of so-called *Szemerédi-Trotter sets* by Shkredov [101]. Shkredov's version of this theorem uses the notation $d_+(A)$, which we omit. The key content of this theorem is that if every element of a set has a lot of representations in a multiplicative sense, then has very little additive structure. The bounds are (increasingly complicated) consequences of the Szemerédi-Trotter theorem.

Theorem 7.12. *Suppose that $A \subseteq \mathbb{R} \setminus \{0\}$ is a finite set, to which we associate two other sets $Q, R \subseteq \mathbb{R}$ with the property that $|Q| \geq \max(|A|, |R|)$ and*

$$r_{QR}(a) \geq t \quad \forall\ a \in A$$

for some $t \geq 1$.

Then we have the following estimates:

(i) $|\{x \in A - A \colon r_{A-A}(x) \geq \tau\}| \ll \frac{|A|^2|Q|^2|R|^2}{t^3\tau^3}$;

(ii) $\mathsf{E}_3(A) \lesssim \frac{|A|^2|Q|^2|R|^2}{t^3}$;

(iii) for any set B, $\mathsf{E}^+(A,B) \ll \frac{|B|^{3/2}|A|^{1/2}|Q||R|}{t^{3/2}}$;

(iv) for any set B *satisfying* $r_{A-A}(b) \geq \Delta$ *for all* $b \in B$*:*

$$\sum_x r_{A-A}^2 r_{B-B}(x) \lesssim \frac{|A|^{17/10}|Q|^{11/5}|R|^{11/5}|B|^{9/10}}{t^{33/10}\Delta^{4/5}}\,.$$

Proof of (i): Let X_τ be the set of $x \in A - A$ satisfying $r_{A-A}(x) \geq \tau$. Then $\tau|X_\tau| \leq |\{d = x - y \colon d \in X_\tau, x, y \in A\}|$; using the Szemerédi-Trotter theorem we have

$$t\tau|X_\tau| \leq |\{d = qr - y\}| \ll (|X_\tau||A||Q||R|)^{2/3} + |Q||A| + |R||X_\tau|\,.$$

If the third term dominates, then $|X_\tau| \geq |A|^2|Q|^2|R|{-}1$; since $|A|^2 \geq |X_\tau|$, this is a contradiction. We rearrange the above to find that

$$|X_\tau| \ll \frac{|A|^2|Q|^2|R|^2}{\tau^3 t^3} + \frac{|Q||A|}{t\tau};$$

if the second term is greater than the first term, then $t\tau > |R|\sqrt{|A||Q|}$. Since $\tau \leq |A|$ and $t \leq R$, we have a contradiction.

Proof of (ii): By a dyadic decomposition and (i), we calculate

$$\begin{aligned}\mathsf{E}_3(A) &\leq \sum_{i=0}^{\log(|A|)} (2^{i+1})^3 |\{x \in A - A : 2^i \leq r_{A-A}(x) \leq 2^{i+1}\}| \\ &\ll \sum_{i=0}^{\log|A|} 2^{3i} \frac{|A|^2|Q|^2|R|^2}{t^3 2^{3i}} \sim \log(|A|)\frac{|A|^2|Q|^2|R|^2}{t^3}\,.\end{aligned}$$

Proof of (iii): As before, letting $X_\tau = \{d \in A - B : r_{A-B} \geq \tau\}$ we estimate

$$\tau t|X_\tau| \leq |\{d = qr - b : q \in Q, r \in R, b \in B, d \in \tilde{X}_\tau\}|\,.$$

We then rearrange to find $|X_\tau| \ll \frac{|B|^2|Q|^2|R|^2}{t^3\tau^3}$ using the same arguments as in (i). Using the estimate

$$\mathsf{E}(A,B) \leq \sum_{x : r_{A-B}(x) \leq \tau} \tau^2 + \sum_{x : r_{A-B}(x) \leq \tau} r_{A-B}^2(x)$$

we find that, after a standard dyadic pigeonholing argument,

$$\mathsf{E}(A,B) \leq \tau |A||B| + \frac{|B|^2|Q|^2|R|^2}{t^3\tau}\,.$$

We optimise by setting $\tau = \frac{|B|^{\frac{1}{2}}|Q||R|}{|A|^{\frac{1}{2}}t^{3/2}}$.

Proof of (iv) We wish to bound the quantity $\sum_x r^2_{A\pm A}(x) r_{B-B}(x)$. By a dyadic pigeonhole argument, there is a set $\tilde{X} \subseteq (A-A)\cap(B-B)$ and a number τ so that $r_{A-A}(\alpha) \sim \tau$ for all $\alpha \in \tilde{X}$. By (i), $|\tilde{X}| \ll \frac{|A|^2|Q|^2|R|^2}{t^3\tau^3}$. Then, on the one hand, we have

$$\sum_x r^2_{A\pm A}(x) r_{B-B}(x) \ll \mathsf{E}(A,B)\tau\,.$$

On the other hand,

$$\begin{aligned}\sum_x r^2_{A\pm A}(x) r_{B-B}(x) &\lesssim \tau^2 \sum_{x\in\tilde{X}} r_{B-B}(x)\\ &\leq \tau^2\Delta^{-1}|\{(x,a_1,a_2,b)\in \tilde{X}\times A\times A\times B : x-a_1 = b-a_2\}|\\ &\leq \tau^2\Delta^{-1}\sqrt{\mathsf{E}(\tilde{X},A)\mathsf{E}(B,A)}\,.\end{aligned}$$

Combining these two estimates with the statement (iii) and the estimate on $|\tilde{X}|$ we find that

$$\begin{aligned}\sum_x r^2_{A\pm A}(x) r_{B-B}(x) &\lesssim \mathsf{E}(A,B)^{1/2}\left(\tau \mathsf{E}(A,B)^{1/2} + \frac{\tau^2\mathsf{E}(\tilde{X},A)^{1/2}}{\Delta}\right)\\ &\ll \mathsf{E}(A,B)^{1/2}\left(\tau \mathsf{E}^(A,B)^{1/2} + \frac{|A|^{7/4}|Q|^2|R|^2}{\Delta\tau^{1/4}t^3}\right)\,.\end{aligned}$$

We optimise, choosing

$$\tau = \frac{|A|^{7/5}|Q|^{8/5}|R|^{8/5}}{\Delta^{4/5}t^{12/5}\mathsf{E}(A,B)^{2/5}}$$

and so

$$\begin{aligned}\sum_x r^2_{A\pm A}(x) r_{B-B}(x) &\lesssim \mathsf{E}(A,B)^{3/5}\frac{|A|^{7/5}|Q|^{8/5}|R|^{8/5}}{\Delta^{4/5}t^{12/5}}\\ &\ll \left(\frac{|B|^{3/2}|A|^{1/2}|Q||R|}{t^{3/2}}\right)^{3/5}\frac{|A|^{7/5}|Q|^{8/5}|R|^{8/5}}{\Delta^{4/5}t^{12/5}}\\ &= \frac{|B|^{9/10}|A|^{17/10}|Q|^{11/5}|R|^{11/5}}{\Delta^{4/5}t^{33/10}}\,.\end{aligned}$$

This proof encompasses [101, Definition 5, Lemma 7 and Lemma 10] of Shkredov.

Proof of Theorem 7.10

The goal of Theorem 7.10 is to find a lower bound on $|A + A|$ for a set $A \subseteq \mathbb{R} \setminus \{0\}$ under the assumption that, for all $a \in A$, there exist sets $Q, R \subseteq \mathbb{R}$ and an integer $t > 0$ such that $r_{QR}(a) \geq t$.

Let $P := \{s \in A + A \colon r_{A+A}(s) \geq \frac{|A|^2}{2|A+A|}\}$ be the set of popular sums of A. These sums support most of the mass of A:

$$|A|^2 = \sum_{x \in A+A} r_{A+A}(x) \leq 2 \sum_{p \in P} r_{A+A}(p)\,.$$

In particular, note that by the pigeonhole principle, there is an element $a_0 \in A$ so that $|(P - a_0) \cap A| \geq |A|/2$.

To prove the statement of the theorem, we use an intermediate quantity X:

$$X := \sum_{x,y,z \in A} r_{A-A}(x - y) P(x + z) P(y + z),$$

where $P(\cdot)$ is the indicator function of P.

An upper bound on X

The upper bound on X is an application of the Cauchy-Schwarz inequality:

$$\begin{aligned} X^2 &\leq \sum_{\alpha,\beta} r_{A+A}^2(\alpha - \beta) P(\alpha) P(\beta) \sum_{\alpha,\beta} \left(\sum_z A(z) A(\alpha - z) A(\beta - z) \right)^2 \\ &= \sum_x r_{A+A}^2(x) r_{P-P}(x) \sum_{\alpha,\beta} \sum_{z,z'} \sum_{a,a',b,b' \in A} \mathbb{1}_{\alpha = a+z = a'+z'} \mathbb{1}_{\beta = b+z = b'+z'} \\ &= \sum_x r_{A+A}^2(x) r_{P-P}(x) |\{((z, z', a, a', b, b') \in A^6 \colon z - z' = a' - a = b' - b\}| \\ &= \sum_x r_{A+A}^2(x) r_{P-P}(x) \mathsf{E}_3(A) \end{aligned}$$

A lower bound on X

The lower bound on the quantity X is somewhat more involved and exemplifies the *eigenvalue method*, introduced by Schoen and Shkredov [96] and Shkredov [102] (see also [72, Lemma 10]).

Choosing appropriate matrices, we realise X as the trace of a product of matrices and use linear algebra to bound the trace.

Let us define two real $|A| \times |A|$ matrices:

$$\mathscr{P} = \Big(P(a+b)\Big)_{a,b} \quad \text{and} \quad \mathfrak{M} = \Big(r_{A-A}(a-b)\Big)_{a,b}.$$

The trace of the product $\mathscr{P}^2\mathfrak{M}$ is precisely the quantity X.

Observe that $\mathscr{P}$ and $\mathfrak{M}$ are symmetric matrices and therefore each have $|A|$ (not necessarily distinct) eigenvalues and $|A|$ corresponding mutually orthonormal eigenvectors [94, Theorem 7.4.6]. In fact, $\mathfrak{M}$ is positive semi-definite – for any $\vec{z} \in \mathbb{R}^{|A|}$, $\vec{z} \cdot \mathfrak{M}\vec{z} \geq 0$. Indeed, indexing $\vec{z}$ by elements of A, we have

$$\vec{z} \cdot \mathfrak{M}\vec{z} = \sum_{a,b \in A} r_{A-A}(a-b) z_a z_b = \sum_{d \in A-A} \left(\sum_{a \in A} A(d-a) z_a \right)^2 \geq 0\,.$$

Let $\mu_1 \geq \mu_2 \geq \ldots \mu_{|A|}$ be the (real) eigenvalues of $\mathscr{P}$ with corresponding normalised eigenvectors $\vec{v}_i$ (so that $\sum_{1 \leq i \leq |A|} v_i^2 = 1$).

We have $\mu_1 \geq |A|/2$. This follows since

$$\mu_1 = \sup_{\vec{x} \in \mathbb{R}^{|A|} : \|\vec{x}\| = 1} \langle \vec{x}, \mathscr{P}\vec{x} \rangle \geq \langle \vec{u}, \mathscr{P}\vec{u} \rangle\,,$$

where $\vec{u} = \frac{1}{\sqrt{|A|}}(1, \ldots, 1)$, and recalling that P is popular *by mass*.

We diagonalise $\mathscr{P}$ and use elementary properties of the trace as well as the positive semi-definiteness of $\mathfrak{M}$ to obtain the inequality:

$$\mathsf{Tr}(\mathscr{P}^2\mathfrak{M}) \geq \mu_1^2 \langle \vec{v}_1, \mathfrak{M}\vec{v}_1 \rangle\,.$$

We claim that $\langle \vec{v}_1, \mathfrak{M}\vec{v}_1 \rangle \geq \frac{|A|^3}{2|A+A|}$. This is the content of Shkredov's [101, Lemma 9], whose proof is contained in [103, Corollary 4.12].

We provide a proof here. Let $\mu_1 = \mu$ denote the corresponding biggest eigenvalue of $\mathscr{P}$. Let $\vec{v}_1 = \vec{v}$ and we index $\vec{v}$ by elements of A. By the Perron-Frobenius theorem, we may assume that all components of $\vec{v}$ are positive.

Observe that v_a, the a-th coordinate of the vector $\vec{v}$, satisfies the equation

$$\mu v_a = \sum_{b \in A} P(a+b) v_b = \sum_x P(x) \sum_a v_{x-a} A(x-a) A(a)\,.$$

Then on the one hand, by an application of Cauchy-Schwarz

$$\begin{aligned}\left(\sum_{a\in A}\mu v_a\right)^2 &= \left(\sum_x P(x)\sum_a v_{x-a}A(x-a)A(a)\right)^2\\ &\leq |P|\sum_x\sum_{a,b\in A} v_{x-a}v_{x-b}A(x-a)A(x-b)\\ &= |P|\sum_{c,d\in A} v_c v_d r_{A-A}(c-d) \leq |A+A|\,\langle \vec{v}, \mathfrak{M}\vec{v}\rangle\,.\end{aligned}$$

Note that here we use that each component of $\vec{v}$ is positive.

On the other hand, we have a lower bound on $\sum_a \mu v_a$ as follows. Let J denote the $|A|\times|A|$ matrix of all ones, and let $\vec{w} = (\sum_a v_a, \ldots, \sum_a v_a)$ be a vector of length $|A|$. Then

$$\left(\sum_a v_a\right)^2 = \vec{v}\cdot\vec{w} = \vec{v}\cdot J\vec{v} \geq \vec{v}\cdot P\vec{v} = \mu\vec{v}\cdot\vec{v} = \mu\sum_a v_a^2 = \mu\,.$$

Again, we use the fact that each component of $\vec{v}$ is positive to realise the equality $\vec{w} = J\vec{v}$.

Hence we have the lower bound $X \geq \mu_1^2\frac{\mu_1^3}{|A+A|} \gg \frac{|A|^5}{|A+A|}$.

Conclusion of the proof of Theorem 7.10

Finally, we combine the upper and lower bounds on X to obtain

$$\begin{aligned}\frac{|A|^{10}}{|A+A|^2} \ll X^2 &\ll \mathsf{E}_3(A)\sum_x r^2_{A+A}(x)r_{P-P}(x)\\ &\lesssim \frac{|A|^2|Q|^2|R|^2}{t^3}\,\frac{|A|^{17/10}|Q|^{11/5}|R|^{11/5}|P|^{9/10}}{t^{33/10}}\left(\frac{|A|^2}{2|A+A|}\right)^{-4/5}\\ &\ll \frac{|A|^{21/10}|Q|^{42/10}|R|^{42/10}|A+A|^{17/10}}{t^{63/10}}\,.\end{aligned}$$

Rearranging completes the proof of Theorem 7.10.

8

Pinned distances in positive characteristic

8.1 Introduction

Suppose we have a finite set of points A in the plane, and we measure the distance between every pair of points. We have made $\binom{|A|}{2}$ measurements, but how many distinct distances are there?

This is the *distinct distances* problem, first asked in 1946, by Erdős [29]. It is easily possible for all $\binom{|A|}{2}$ distances to be distinct: a random set would achieve this [1] The challenge of the distinct distances problem is to find a *lower* bound: given a point set $\mathcal{P}$, how many distinct distances are we guaranteed to have?

If a point set A determines very few distinct distances, then many of the $\binom{|A|}{2}$ measurements between pairs of points of A must coincide. In other words, the set A must be very structured. With this concept in mind, Erdős turned to the most natural example of a structured set of points in the plane: for N a square, the grid of points $\{1, \dots, \sqrt{N}\} \times \{1 \dots, \sqrt{N}\}$. Erdős conjectured that this example provides an asymptotically optimal lower bound:

Conjecture 8.1 (Erdős [29])**.** *There exists $c > 0$ such that for any set $A \subseteq \mathbb{R}^2$ of N points, the number of distinct distances between pairs of points in A is at least $c\frac{N}{\sqrt{\log N}}$.*

[1] For the reader preferring a more explicit example, there are $\binom{|A|}{2}$ distances determined by the set of points $A = \{(1, 2^i) : 1 \leq i \leq N\}$.

A recent breakthrough result of Guth and Katz [43] confirmed Erdős' conjecture, up to a logarithmic factor.

Theorem 8.2 (Guth – Katz [43])**.** *Let $c > 0$ be an absolute constant, and let $A \subseteq \mathbb{R}^2$ be a set of N points. Then A determines at least $c\frac{N}{\log N}$ distances.*

Guth and Katz's result was based upon the Elekes-Sharir framework [28], in which the distinct distance problem was reduced to that of an incidence problem in three dimensions. Closing the gap between Theorem 8.2 and Conjecture 8.1 appears to be currently out of reach.

There are two immediate directions for progress: the first is to ask what happens in higher dimensions; the second, which will be the subject of this chapter, is to ask about distances in sets of points over arbitrary fields.

The main result of this chapter is the following theorem:

Theorem 8.3. *Let $A \subseteq \mathbb{F}^2$ be a non-empty set of points with not all points lying on a single isotropic line. If the characteristic of $\mathbb{F}$ is $p > 0$, suppose in addition that $|A| \leq p^{4/3}$.*

Then A determines at least $c|A|^{2/3}$ distances for some absolute constant $c > 0$.

The condition regarding isotropic lines is a technicality which does not appear over $\mathbb{R}$.

Structure of this chapter

We begin with a discussion on why the distinct distance problem over $\mathbb{F}$ is different to the reals. In particular, in finite fields we have the related Erdős-Falconer problem. Section 8.4 reviews the literature surrounding the distinct distances problem, both over the reals and over arbitrary fields.

To prove Theorem 8.3, we actually provide a new estimate for the stronger *pinned distance* problem. Section 8.3 introduces this. The subsequent chapters are devoted to proving Theorem 8.3.

8.2 Distinct distances over $\mathbb{F}$

What changes?

In the two-dimensional vector space $\mathbb{F}^2$ there are a number of obstructions to the distinct distance problem that do not arise over $\mathbb{R}$.

The first is a somewhat trivial issue: when we take the Euclidean distance between points a and b over the reals, we implicitly take the positive square root of $(a-b)\cdot(a-b)$. We rely on the fact that, since $(a-b)\cdot(a-b)\geq 0$, taking the square root is a well-defined operation. This is no longer true over arbitrary fields, where the concept of 'positive' is undefined. To get around this issue, we use a (standard) modified definition of distance.

Definition 8.4. *The distance d between two points $x=(x_1,x_2)$, $y=(y_1,y_2)$ in the plane $\mathbb{F}^2$ is*

$$d(x,y) := (x-y)\cdot(x-y) = (x_1-y_1)^2+(x_2-y_2)^2\,.$$

We are interested in a lower bound on number of distinct distances that can arise between pairs of points in A.

Definition 8.5. *For a set $A\subseteq\mathbb{F}^2$ the set of distances determined by A is $\{d(a,b)\colon a,b\in A, a\neq b\}$.*

The number of distinct distances determined by A is

$$\Delta(A) := |\{d(a,b)\colon a,b\in A \text{ and } a\neq b\}|\,.$$

A less trivial obstruction, which does not appear over the real numbers, is that of isotropy: suppose, in $\mathbb{R}$ that we have points x,y such that $d(x,y)=0$. It follows automatically from the definition of the inner product that $x=y$. However, suppose now that $x,y\in\mathbb{C}$, with $x=(1,i)$ and $y=(0,0)$. Then $d(x,y)=(1-0)^2+(i-0)^2=0$, but the points x and y are not the same. In fact, any two points lying on the line passing through x and y will be distance 0 apart.

In particular, if N points in $\mathbb{C}$ lie on this ('isotropic') line, then they can only determine one distinct distance, namely the distance 0. Given that we want to find a lower bound on the number of distinct distances, this is indeed a problem. We cannot avoid this type of example, which is present in any field in which -1 is a square, and so must place a condition to ensure that it is not the case that all N points are distance 0 apart.

A third technical obstruction to the problem is one unique to *finite* fields. In (the infinite field) $\mathbb{R}$, it is always reasonable to ask for a lower bound on the number of distinct distances determined by N points in terms of N. After all, there is no reason for the number of distances to stop growing once N reaches a certain threshold. In a finite field however, this is not the case. For example, N points in $\mathbb{F}_p^2$ can never determine more than $|\mathbb{F}_p| = p$ distances. To this end, we must place a restriction on N. From the above example of $\mathbb{F} = \mathbb{F}_p$, it is tempting to suppose that restriction in terms of the cardinality of the field would suffice. However, we must also take into account subfield-interaction, as the following example demonstrates. The

Recall that we wish to find a lower bound on the number of distinct distances determined by N points; we want this lower bound to be given in terms of N. In the real setting, the set of attainable distances cannot be larger than the 'cardinality' of the (infinite) field – i.e the distance function restricted to finite sets is not surjective. In a finite field $\mathbb{F}_q$ however, the set of possible distances has size q, and so it could well be the case that the set of attainable distances can be (a positive proportion) of this. For this reason, we enforce a bound on the size of the set $|A|$: if $|A|$ is smaller than this threshold bound, then it is sensible to ask for the number of distinct distances attained by A as a function of the cardinality of $|A|$. In deciding this threshold, we require that $|A|$ is actually bounded in terms of the *characteristic* of the field, and not just the cardinality of the field. The reason for this is demonstrated by the following example.

Example 8.6. *Consider the set of points $A = \mathbb{F}_p \times \mathbb{F}_p \subseteq \mathbb{F}_{p^r} \times \mathbb{F}_{p^r}$. Then $|A| = p^2$, and $\Delta(A) \geq p - 1$. That is, A determines at least $p - 1$ distances; in particular for all $r \in \mathbb{N}$ the set of distances is $\mathbb{F}_p$.*

As we take $r \to \infty$, the set A will occupy an increasingly minuscule proportion of the field. That is, $\lim_{r\to\infty} \Delta(A)/|\mathbb{F}_{p^r}| = \lim_{r\to\infty} p^{1-r} = 0$. The obstruction in this case is that A is a Cartesian product of subfields of the field.

In this example we see that the threshold at which A determines a positive proportion of attainable distinct distances depends upon the interaction of A with subfields of its ground field.

The case of large sets

As demonstrated by taking $A = \mathbb{F}_q^2$, in the finite field setting if the set $A \subseteq \mathbb{F}_q^2$ is sufficiently large relative to q, then we expect many (in terms of the characteristic of the field) distances to occur. Determining the threshold of 'sufficiently large' is known as the Erdős-Falconer problem, named after a measure-theoretic analogue by Falconer [32] to the distance problem.

This type of viewpoint naturally studies large sets, and typically uses Fourier-analytic techniques. In contrast, the line of reasoning pursued in this chapter will aim to achieve results pertaining to small sets – results of the type "given a set A, how many distances are we guaranteed to find?". In this regime, Fourier techniques are unsuitable, and instead new results have relied on incidence theorems.

The first Erdős-Falconer type result was by Iosevich and Rudnev [49], valid in arbitrary dimension.

Theorem 8.7 (Iosevich-Rudnev [49])**.** *For all $d \geq 2$ and odd q, if $A \subseteq \mathbb{F}_q^d$ has cardinality $|A| > 4q^{\frac{d+1}{2}}$, then $\Delta(A) = q$.*

In the two dimensional case, Iosevich and Rudnev's exponent of $3/2$ was improved to the smaller exponent of $4/3$. That is, whenever $|A| \gg q^{4/3}$ then $\Delta(A) \gg q$. This is a result of Chapman et al. [19, Theorem 2.2], who used Fourier analysis to show that for any $|A| \gg q^{4/3}$, $\Delta(A) \gg q$. The latter paper claimed the bound for $q \equiv 3 \bmod 4$ only; it was then observed by Bennett et al. [8, Theorem 1.6] that the same proof works for $q \equiv 1 \bmod 4$ as well.

Notice that Iosevich and Rudnev's result, although weaker in terms of the exponent, guarantees that A determines *all* distances, whereas the results proving the threshold $4/3$ can guarantee only a positive proportion of distances.

In fact, Murphy and Petridis [69] show that this 'positive-proportion' type result is the best that can be achieved, where we think of q in the preceding discussion as $q = p^r$:

Theorem 8.8 (Murphy, Petridis [69])**.** *Let p be a prime. Then there exist infinitely many $r \in \mathbb{N}$ and $A \subseteq \mathbb{F}_{p^r}^2$ such that $|A| = (p^r)^{4/3}$ and $\Delta(A) \neq p^r$. In fact, there exists $r_0 = r_0(p)$ so that for all $r \geq r_0$, we have the bound $\Delta(A) \leq p^r/2$.*

The main study of the two-dimensional Erdős-Falconer problem over $\mathbb{F}_q$ is now to obtain the correct value for the constant. That is, if we have $|A| = q^{4/3}$,

what is the smallest value of $c > 0$ so that $\Delta(A) \geq cq$. If we write $q = p^r$, it is of interest to determine how c depends on both q and r. Of course, there remains much to understand in other cases, particularly in higher dimensions or if the set of points in question is endowed with a particular structure. We refer the reader to Koh, Pham and Vinh [56] and references therein for recent developments in this direction.

For the remainder of this chapter, we will think of the set of points as being small with respect to the characteristic of the field. That is, the cardinality of the point set will be bounded by a sub-quadratic function of the characteristic of the field. Note that this does not address the case when the number of points in a finite field is (quantifiably) *large* in terms of the characteristic of the field, but *small* in terms of the cardinality of the field.

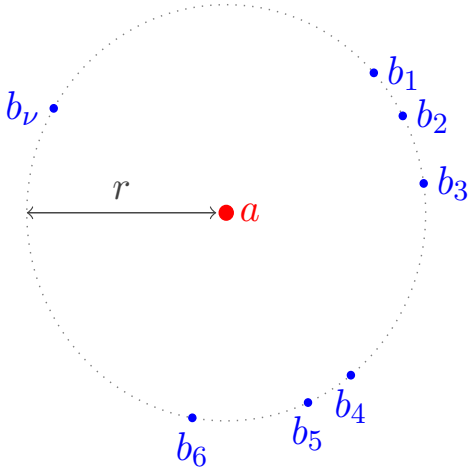
Trivial distance bound

It is instructive to consider first the trivial bound on the number of distances of a set $A \subseteq \mathbb{F} \times \mathbb{F}$, first observed by Erdős [29] over the reals. Assuming that -1 is not a square (and so we cannot fall into the isotropic case), the trivial distance bound is that a set A of N points in $\mathbb{F} \times \mathbb{F}$ determines at least $\sqrt{(N-1)/2}$ distinct distances.

Indeed, if, on the one hand, a distance pinned at the point $a \in A$ repeats at most ν times, for some parameter ν, then there are at least $(N-1)/\nu$ distances. Suppose on the other hand, that a distance r from the point a repeats at least ν times.

Then there are ν points, say $b_1, \ldots, b_\nu$ on the circle of radius r centred at a.

Figure 8.1: If a distance r from a repeats with multiplicity ν, then ν points lie on the circle of radius r centred at a



Consider the distances between b_1 and $b_2, \ldots, b_\nu$: if a distances is repeated, say $d(b_1, b_i) = d(b_1, b_j)$, then b_i, b_j lie on a circle centred at b_1. The points b_i and b_j also lie on a circle centred at a, hence are the intersection points of these two (distinct) circles. This means that there can be at most two such points b_i, b_j, and so there are at least $(\nu - 1)/2$ distances pinned at b_1. Thus $\Delta(A) \geq \max(N/\nu, (\nu - 1)/2)$; optimisation of the parameter ν yields the result.

In fact, this proof shows a stronger result than the existence of $N^{1/2}$ distances: it shows that there exists a *special* point from which many distances occur. In the notation of the example, either a or b_1 is this special point. This stronger variant is called the *pinned distance problem*.

8.3 The pinned distance problem

The pinned distance problem is to show that any set of points A contains a point $a_0 \in A$ such that the cardinality of the set $\Delta_{a_0}(A) := \{d(a_0, b) \colon b \in A\}$ is large. It is clear that $\Delta(A) \geq |\{d(a_0, b) \colon b \in A\}|$.

We write

$$\Delta_{\text{pin}}(A) = \max_{a \in A} |\Delta_a(A)| \,.$$

The pinned distance problem is to find the correct minimum lower bound for $\Delta_{\text{pin}}(A)$ in terms of $|A|$ over all sets A.

In the real setting, Erdős [30] conjectured in 1975 that asymptotically, the lower bound for $\Delta_{\text{pin}}(A)$ should match that of $\Delta(A)$. In fact, he conjectured

the more global statement $\sum_{a \in A} |\Delta_a(A)| \geq c|A|^2(\log |A|)^{-1/2}$ for some $c > 0$, which is stronger than the original distinct distances conjecture. This global statement should be interpreted as a quantitative version of the statement 'many distances are determined from almost all points A'.

Whilst the distinct distances bound has been (almost) resolved in $\mathbb{R}^2$ [43], the pinned distance variant remains open. We take a brief digression and review the literature of these problems, both in the reals, and in the arbitrary field setting which we shall study in this chapter.

8.4 Literature review

The distinct distances problem over $\mathbb{R}$

After Erdős [29] proved the first pinned distance result, the next pinned distance records on sets of n points in the plane were the result of involved elementary geometric arguments (e.g. Moser [67], Chung[21]).

The first, as far as the author is aware, pinned distance result also marks the introduction of incidence geometry to the distinct distances problem. Clarkson, Edelsbrunner, Guibas, Sharir and Welzl [23] proved that in any set of n points in $\mathbb{R}^2$, there exists a point determining at least $n^{3/4}$ pinned distances. This was a consequence of an incidence bound between points and spheres.

Chung, Szemerédi and Trotter [22] then brought the Szemerédi-Trotter theorem to the problem[2]. Chung et al. proved that n points determine at least $n^{4/5-\epsilon}$ pinned distances.

Using graph-theoretic techniques, namely the Crossing Lemma, Székely [109] proved that n points always determine $\Omega(n^{4/5})$ pinned distances. Székely's result is actually a pinned distance result: he proves that there exists a point determining this many distances.

Further progress on the pinned distance variant stemmed from a technique of Solymosi and Tóth [107] using Beck's theorem and the Szemerédi-Trotter theorem.

Finally, a series of papers by Tardos [113], Katz [53] and Tardos and Katz [55] improved this exponent to the current record. Using an entropy method, Katz and Tardos actually prove a lower bound on the number of distinct

[2]Missing here is perhaps a result of Beck referenced within other works, whose preprint entitled 'Different distances' I have been unable to track down.

pairwise sums formed by adding elements of the same row of an $n \times s$ real matrix; Solymosi and Toth [107] (implicitly) demonstrated the connection between this and the distinct distances problem; Tardos [113] explicitly shows this connection. Katz and Tardos [55] prove the following:

$$\Delta_{\text{pin}}(A) \gtrsim |A|^{\frac{48-14e}{55-16e}} \sim |A|^{0.8641\ldots}.$$

The distinct distance problem however has enjoyed much more progress. It was solved (up to a logarithmic factor) by Guth and Katz [43], as stated in Theorem 8.2. In fact, this is the first *non*-pinned distance result. We will follow a related approach, and so we discuss their method and the Elekes-Sharir framework [28] in this section.

The Elekes–Sharir Framework

Elekes and Sharir's [28] key innovation was to realise the distance problem as problem of counting equivalence classes of points modulo rigid motions. Informally, the set of rigid motions can be thought of as the set of translations and rotations of a point set.

Counting a distance which repeats in a point set $A \subseteq \mathbb{R}^2$ at least k times is related to counting the number of rigid motions g for which $|A \cap gA| \geq k$. Elekes and Sharir then reduced this to an incidence problem between points and helices in three dimensions. Their framework was the inspiration for the breakthrough work of Guth and Katz.

Guth and Katz realised the problem of counting rigid motions as an incidence bound in three-space, not between points and helices, but between points and lines. This straightening they use can be interpreted as an instance of the Blaschke-Grünwald kinematic mapping, which is discussed in Section 8.7. They then solved this ensuing incidence bound, enabling their (almost) tight bound.

Related work over $\mathbb{F}$

Over $\mathbb{F}$ the state of affairs is much different.

As previously mentioned, the first non-trivial result in a field other than $\mathbb{R}$ was by Bourgain, Katz and Tao [17], and was a consequence of their non-trivial incidence bound between points and lines. Their result is valid for $\mathbb{F} = \mathbb{F}_p$; the strongest instance of this result is an incidence of the Stevens-de Zeeuw [108]

incidence bound valid for arbitrary $\mathbb{F}$, presented in Chapter 5 as Theorem 5.4. This technique actually elicits a pinned distance result.

A recent strengthening of this technique by Iosevich et al [48, 47] has recently improved the exponent to $\frac{1}{2} + \frac{149}{4215}$ by using bounds on the additive energy of a set lying on a paraboloid. This result is valid for $\mathbb{F} = \mathbb{F}_p$ and $p \equiv 3 \bmod 4$. This latter quantity is estimated and bounded within the restriction theory literature – again, this is a consequence of Lewko's application of incidence geometry to this problem [61].

A powerful cause of progress for the pinned distance problem is of a geometric nature, involving studying perpendicular bisectors. This began by Hanson, Lund and Roche-Newton [44] as an analogue for a concurrent work of Lund, Sheffer and de Zeeuw [64]. This work is an Erdős-Falconer pinned distance variant, matching Chapman's [19] exponent.

The most recent progress on the distance problem is by Lund and Petridis [63], which states that, under suitable conditions on a set of N points in terms of the cardinality of the field, that N (non-isotropic) points in the plane determine at least $N^{20/37}$ distances. Lund and Petridis proved a pinned distance result; it is this technique that is most similar to our methods.

Theorem 8.9 (Lund-Petridis [63])**.** *Let $\mathbb{F}$ be a field, and $A \subseteq \mathbb{F} \times \mathbb{F}$ with not all points of A lying on an isotropic line.*
If the characteristic of $\mathbb{F}$ is $p > 0$, suppose also that $|A| \leq p^{8/5}$.

Then there exists $a \in A$ such that $|\{d(a,b) : b \in A\}| \gg |A|^{20/37}$, and in particular, $\Delta(A) \gg |A|^{20/37}$.

In the case of the point set being a Cartesian product, and $\mathbb{F} = \mathbb{F}_p$ Petridis [81] has proved a stronger result: if $A = X \times X \subseteq \mathbb{F} \times \mathbb{F}$, then $\Delta_{\text{pin}}(A) \gg \min\{p, |A|^{3/2}\}$.

8.5 Main Results

We prove a lower bound on the number of pinned distances $\Delta_{\text{pin}}(A)$. The main result of this chapter is the following theorem, of which Theorem 8.3 is an immediate consequence.

Theorem 8.10. *Let $A \subset \mathbb{F}^2$ be a set of points, not all lying on a single isotropic line. If the characteristic of $\mathbb{F}$ is $p > 0$, suppose also that $|A| \leq p^{4/3}$.*

Then

$$\Delta_{pin}(A) \gg |A|^{2/3}\,. \tag{8.1}$$

Sub-optimal Result

Perhaps unusually, we choose to present a previous version of Theorem 8.10. This version is quantitatively weaker in all senses: both the exponent and the range of validity are weaker. The reason for presenting this sub-optimal result is because of an interesting 'complexification' technique, whereby an object in $\mathbb{F}_p \times \mathbb{F}_p$ is analysed by mapping it to $\mathbb{F}_{p^2}$. A direct analogy is that we interpret $(x, y) \in \mathbb{R} \times \mathbb{R}$ as $x + iy \in \mathbb{C}$. This analysis is ultimately redundant in this situation, but the technique remains of interest and may find applications in the future.

Theorem 8.11 (Suboptimal version of Theorem 8.10)**.** *Let $A \subset \mathbb{F}^2$ be a set of points, not all lying on a single isotropic line. If the characteristic of $\mathbb{F}$ is $p > 0$, suppose also that $|A| \leq p^{4/3}$. Then*

$$\Delta_{pin}(A) \gg |A|^{5/8}\,. \tag{8.2}$$

In the special case where $\mathbb{F} = \mathbb{F}_p$, $p \equiv 3 \bmod 4$ and $|A| \leq p^{10/17}$, one has

$$\Delta_{pin}(A) \gg |A|^{13/20}\,. \tag{8.3}$$

8.6 Discussion of techniques

High-level overview of pinned distances strategy

For a fixed a_0, the set of all distances pinned at a_0 is large *unless* there are many configurations containing points $(x, y) \in A^2$ satisfying $d(a_0, x) = d(a_0, y)$.

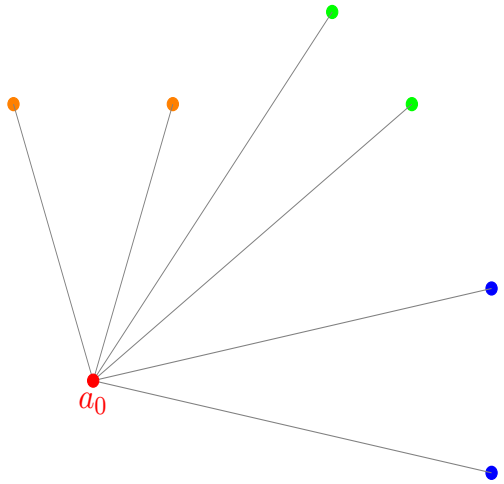


Figure 8.2: a_0 determines many distances unless there are many triangles

If this happens, then there are lots of isosceles triangles (whose 'apex' is at the pinned point a_0). We will show (a quantitative version of the qualitative statement) that there cannot be too many such isosceles triangles, unless we are in a particularly structured regime that already gives many pinned distances.

Pinned distance strategy: perpendicular bisectors

To prove Theorems 8.10 and 8.11 we study the set of perpendicular bisectors defined by pairs of points in A. The perpendicular bisector of $a, b \in \mathbb{F}^2$ with $d(a,b) \neq 0$ is defined as the line

$$\mathsf{B}(a,b) = \{x \in \mathbb{F}^2 \colon d(a,x) = d(b,x)\}\,.$$

To provide some intuition to the relevance of perpendicular bisectors to the pinned distance problem, consider the set of distances of A pinned at the origin $(0,0)$. If A is a 'minimal configuration' – that is, A determines few distinct distances – then there are many repeated distances. In particular, there is a pair $(a,b) \in A \times A$ such that $d(0,a) = d(0,b)$. Notice that the origin, the pin in question, lies on the perpendicular bisector of a and b.

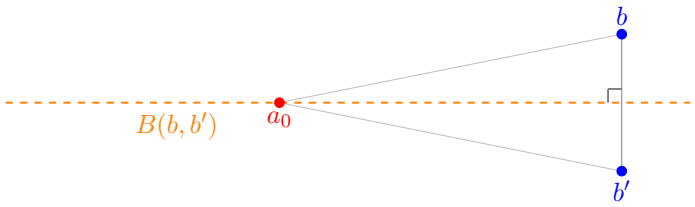


Figure 8.3: The pin lies on the perpendicular bisector

We prove Theorem 8.11 by considering (a subtle variant of) the *bisector energy* of the set A. The bisector energy of the set A counts pairs of points in A whose perpendicular bisector coincides:

$$|\{(b, b', c, c') \in A^4 : \mathsf{B}(b, b') = \mathsf{B}(c, c')\}| .$$

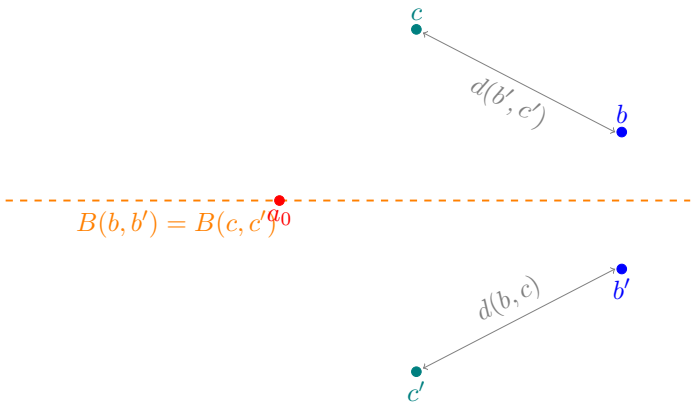


Figure 8.4: The bisector energy counts quadruples of points

Our variant of the bisector energy – $\sum_\ell n_2(\ell)^2$ in Section 8.8 – has a further technical condition allowing us to disregard the delicacies that arise from isotropic lines.

Over the reals, Lund and Petridis show [63, Theorem 2] (a quantified version of the statement) that if the bisector energy is large, then A contains many collinear points or many co-circular points. In our proof, this statement naturally arises.

The number of isosceles triangles in A is controlled by the bisector energy. We arrive at the dichotomy that either the bisector energy is small (and thus the number of isosceles triangles is small), or the set A has structure (in the sense that A contains many co-circular or co-linear points).

The new idea that we use to prove Theorems 8.10 and 8.11 is to realise the bisector energy of n segments of the same length as an incidence problem between n points and n planes. We use the (general field analogue of the) *kinematic mapping* of Blaschke and Grünwald [11, 42], which embeds the space of segments of the same length into projective three-space. We then use Rudnev's point-plane theorem (Theorem 2.9).

Counting perpendicular bisectors with incidences

To be precise, let $S_r = S_r(A) := \{(a,b) \in A^2 : d(a,b) = r\}$ be the set of pairs of points distance r apart. The distance problem is first transformed into a question on isosceles triangles. This is motivated by the idea that if a pinned point has many repeated distances, then there are many isosceles triangles with the pin as their 'apex'. The number of isosceles triangles is then related to an incidence problem between $|S_r(A)|$ points and $|S_r(A)|$ planes.

In the proof, we use an *algebraically closed* field $\mathbb{F} = \bar{\mathbb{F}}$. Theorems 8.11 and 8.10 as stated (where $\mathbb{F}$ is not required to be algebraically closed) follow since $\mathbb{F} \subset \bar{\mathbb{F}}$, where $\bar{\mathbb{F}}$ is the algebraic closure of $\mathbb{F}$; clearly then $|A| = |\bar{A}|$ and $|S_r(A)| = |S_r(\bar{A})|$.

It then remains to estimate $\sum_{r>0} |S_r(A)|^2$. The bounds in Theorem 8.11 arise from trivially estimating this quantity; the improvement for the case $\mathbb{F} = \mathbb{F}_p$ Theorem 8.11 follows from the following proposition:

Proposition 8.12. *Suppose $\mathbb{F} = \mathbb{F}_p$ and $p \equiv 3 \bmod 4$. For $|A| \leq p^{10/17}$ we have*

$$\sum_{r \neq 0} |S_r|^2 \ll |A|^{3+\frac{2}{5}} . \tag{8.4}$$

To prove Theorem 8.10, we avoid using Proposition 8.12; if we are in a situation where the term involving $\sum |S_r|^2$ dominates, we must already be in a case where we have few triangles, and so we use this observation instead. The proof of Proposition 8.12 uses the Elekes-Sharir paradigm [27]: bounding $\sum |S_r|^2$ is related to bounding k-rich rigid motions.

The finite field analogue of this according incidence bound is a theorem of Kollár [57] (whose aim was to provide an 'algebraic' proof of the Guth-Katz incidence bound). However, this approach is not enough to estimate the number of 'very rich' rigid motions, as Kollár's theorem is trivial in this range. For larger values of k, we use a different approach, estimating the number

of k-rich rigid motions via a complexification argument: rigid motions in $\mathbb{F}_p$ become affine transformations in $\mathbb{F}_{p^2}$. Using a positive characteristic analogue of a technique by Solymosi and Tardos [106, Theorem 3], we reduce this to an incidence problem between points and lines.

However, as previously stated, this argument to bound $\sum |S_r|^2$ is redundant. This is because of a easy rearrangement, which we had overlooked. This was pointed out by both Giorgis Petridis and Thang Pham: we will describe this later in the context of the inequalities in question.

8.7 A toolkit for distinct distances

In this section we develop the necessary mathematical tools to prove Theorems 8.10 and 8.11. We begin by defining the kinematic mapping which will transform rigid motions into projective points. We also introduce the Clifford-algebra framework which we will later use. Isotropy has been mentioned as an obstruction to achieving many distinct distances; in Section 8.7 we formally define this, and define perpendicular bisectors of non-isotropic segments. Finally, in Section 8.7 we introduce axial symmetries; these become projective planes via the kinematic mapping.

A framework for distance preserving transformations

To count the number of distinct distances, we will study the group of distance-preserving transformations of the affine plane $\mathbb{F}^2$. An element of this isometry group is a composition of a translation function and a distance-preserving linear map. We denote the roup of distance-preserving linear maps in $GL_2(\mathbb{F})$ as $O_2(\mathbb{F})$.

We define the isometry group of $\mathbb{F}^2$ as follows:

$$\mathsf{Isom}(\mathbb{F}^2) := \{x \mapsto Mx + t \colon M \in O_2(\mathbb{F}), t \in \mathbb{F}^2\}\,.$$

The orthogonal group $O_2(\mathbb{F})$ can be identified with the set of 2×2 matrices M with elements in $\mathbb{F}$ satisfying $M^\top M = I$. Elements of $O_2(\mathbb{F})$ have unit determinant.

The group $\mathsf{Isom}(\mathbb{F}^2)$ has two important cosets, which is a consequence of the group $O_2(\mathbb{F})$ having two cosets. We will identify an element of $O_2(\mathbb{F})$ with the matrix M, and say that M is a *rotation* or *orientation-preserving* if $\det(M) = 1$ and M is a *reflection* or *orientation-reversing* if $\det(M) = -1$. We

recall that since the characteristic of $\mathbb{F}$ is necessarily different to 2, reflections and rotations are distinct. Over the reals, preserving orientation has an obvious physical interpretation; over $\mathbb{F}$ however, this intuition quickly deteriorates. The coset of rotations is in fact a subgroup of $O_2(\mathbb{F})$, denoted $S0_2(\mathbb{F})$. Explicitly

$$SO_2(\mathbb{F}) := \{g \in SL_2(\mathbb{F}) \colon \forall x, y \in \mathbb{F}^2,\, d(x,y) = d(gx, gy)\}.$$

After identifying distance preserving transformations with matrices, we have

$$SO_2(\mathbb{F}) = \left\{ \begin{pmatrix} u & -v \\ v & u \end{pmatrix} : u, v \in \mathbb{F}, u^2 + v^2 = 1 \right\}.$$

We now return to defining the index-two cosets of $\mathsf{Isom}(\mathbb{F}^2)$. Let $T_2(\mathbb{F})$ be the group of translations $x \mapsto x + t$ acting on the plane $\mathbb{F}^2$.

The subgroup $\mathrm{SF}_2(\mathbb{F})$ of positively-oriented rigid motions of $\mathbb{F}^2$ is generated by $SO_2(\mathbb{F})$ and $T_2(\mathbb{F})$; this is the analogue of the special Euclidean group $\mathrm{SE}_2(\mathbb{R})$, and can be thought of as the set of distance-preserving transformations that are a composition of a rotation and a translation.

The other coset of $\mathsf{Isom}(\mathbb{F}^2)$ is the set of axial symmetries: distance-preserving transformations that are a composition of a non-trivial reflection and a translation. This coset is not a subgroup.

Since we will take advantage of the fact that $\mathrm{SF}_2(\mathbb{F})$ is a group, we further develop the notation associated to this subgroup.

There is an injective group homomorphism from the group $\mathrm{SF}_2(\mathbb{F})$ (where the group operation is composition of maps) into $\mathsf{SL}_3(\mathbb{F})$ (where the group operation is matrix multiplication), defined by:

$$\left(\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} u & -v \\ v & u \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} s \\ t \end{pmatrix} \right) \mapsto \begin{pmatrix} u & -v & s \\ v & u & t \\ 0 & 0 & 1 \end{pmatrix}, \tag{8.5}$$

where $u^2 + v^2 = 1$. This acts projectively on $\mathbb{F}^2$: we identify $(x, y) \in \mathbb{F}^2$ naturally with $[x : y : 1] \in \mathbb{PF}^3$; the element described by (8.5) acts on (x, y) by 'rotating' it according to the values of u, v, and then translating it by (s, t).

Thus, we see that $d(x, y) = d(x', y')$ if and only if there exists $g \in \mathrm{SF}_2(\mathbb{F})$ such that $g(x, y) = (x', y')$. If such a g exists, an easy calculation shows that it is unique.

Blaschke-Grünwald Kinematic Mapping

The Blaschke-Grünwald kinematic mapping [11, 42] assigns to an element $g \in \mathrm{SE}_2(\mathbb{R})$, given by a matrix (8.5) – a composition of a rotation and translation – a projective space point in $\mathbb{PR}^3$. The rotation is about the origin by angle θ where $\cos\theta = u$. Guth and Katz reinvented a special case of this mapping in [43].

For its analogue in arbitrary fields, we will define the mapping over an algebraically closed field $\bar{\mathbb{F}}$. For notational convenience, we let $\mathbb{F} = \bar{\mathbb{F}}$.

Let $\mathcal{C} \subseteq \mathbb{F}^2$ denote the unit circle.

The algebraically closed field means that for all $(u, v) \in \mathcal{C}$, we can find values $(\tilde{u}, \tilde{v}) \in \mathbb{F}^2$ satisfying $\tilde{u}^2 = \frac{1+u}{2}$ such that:

$$u = \tilde{u}^2 - \tilde{v}^2, \qquad v = 2\tilde{u}\tilde{v}\,. \tag{8.6}$$

One can show that $(\tilde{u}, \tilde{v}) \in \mathcal{C}$. The choice of root taken as a solution to the equation $\tilde{u}^2 = \frac{1+u}{2}$ we choose for $\tilde{u}$ will define $\tilde{v}$ uniquely.

We are now ready to define an appropriate reinterpretation of the original Blaschke-Grünwald kinematic mapping $\kappa : \mathrm{SF}_2(\mathbb{F}) \to \mathbb{PF}^3$. Under this mapping, an element of $\mathrm{SF}_2(\mathbb{F})$ of the form of (8.5) becomes the projective point:

$$[X_0 : X_1 : X_2 : X_3] = [2\tilde{u} : 2\tilde{v} : s\tilde{u} + t\tilde{v} : s\tilde{v} - t\tilde{u}]\,. \tag{8.7}$$

Conversely we can recover u, v, s, t from a point in the image of κ:

$$u = \frac{X_0^2 - X_1^2}{X_0^2 + X_1^2},\ v = \frac{2X_0X_1}{X_0^2 + X_1^2},\ \frac{s}{2} = \frac{X_1X_3 + X_0X_2}{X_0^2 + X_1^2},\ \frac{t}{2} = \frac{X_1X_2 - X_0X_3}{X_0^2 + X_1^2}\,. \tag{8.8}$$

Remark 8.13. *We can avoid turning to the algebraic closure, by say multiplying by $\tilde{u} \neq 0$. Substitution of* (8.6) *yields*

$$[X_0 : X_1 : X_2 : X_3] = [u + 1 : v : \frac{s(u+1) + tv}{2} : \frac{sv - t(u+1)}{2}]\,.$$

This fails when $\tilde{u} = 0$; in this case, we have $\tilde{v} \neq 0$, so multiply by $\tilde{v}$ instead.

The image of the kinematic mapping κ, is $\mathbb{PF}^3 \setminus \{X_0^2 + X_1^2 = 0\}$.

We can express the composition of two rigid motions g and h in $\mathrm{SF}_2(\mathbb{F})$ as a multiplication in the Clifford algebra $\mathrm{CL}(0, 2, 1)$: we identify an element $g \in \mathrm{SF}_2(\mathbb{F})$ given by (8.5) with the algebra element

$$g = X_0 + X_1e_1e_2 + X_2e_1e_3 + X_3e_3e_2\,,$$

where e_1, e_2, e_3 are anti-commuting generators such that $e_1^2 = e_2^2 = -1$ and $e_3^2 = 0$.[3]

We can calculate the composition of g and h explicitly: let g be as above, and let

$$h = Y_0 + Y_1 e_1 e_2 + Y_2 e_1 e_3 + Y_3 e_3 e_2 \,.$$

Then we can calculate the product gh as a matrix multiplication using the substitution (8.8) and ordinary multiplication rules. We can calculate their product as a Clifford algebra multiplication by the generator multiplication rules described above. Converting the matrix multiplication to a Clifford algebra via the kinematic map shows that they coincide.

In particular, their product is:

$$[X_0Y_0 - X_1Y_1 : X_1Y_0 + X_0Y_1 : X_0Y_2 + X_1Y_3 + X_2Y_0 - X_3Y_1 : X_0Y_3 - X_1Y_2 + X_3Y_0 + X_2Y_1] \,.$$

The group product $gh \in \mathrm{SF}_2(\mathbb{F})$ corresponds to the product of the two corresponding Clifford algebra elements, which one can verify by a calculation.

In particular, the kinematic mapping image $\kappa(gh)$ of the product $gh \in \mathrm{SF}_2(\mathbb{F})$ is represented by the Clifford algebra product of the Clifford algebra elements, corresponding to $\kappa(g)$ and $\kappa(h)$; this is a computation which we omit.

Isotropic lines and Perpendicular Bisectors

A vector $v \neq 0 \in \mathbb{F}^2$ is *isotropic* if $d(v, v) = 0$.

To be precise, isotropy is defined relative to the symmetric bilinear form with which we assume the vector space in which we are working is equipped. Suppose V is a vector space over $\mathbb{F}$ equipped with a symmetric bilinear form $B : V \times V \to \mathbb{F}$. Then V is an isotropic space if there exists a non-zero $v \in V$ with $B(v, v) = 0$. We could equivalently define V to be isotropic if its associated quadratic form has a non-trivial kernel. With this second definition, it is clear that the set of isotropic vectors (that is, $v \neq 0$ satisfying $B(v, v) = 0$) together with the element 0, forms a subspace. In the context of $V = \mathbb{F} \times \mathbb{F}$, this subspace is a line, which we later refer to as an isotropic line. We mention in passing, but do not use, the fact that if V is a vector space as above of dimension $\dim(V) \geq 3$ and if $\mathbb{F}$ is finite, then V is isotropic (see e.g.

[3]The notation of the Clifford algebra $\mathrm{Cl}(0, 2, 1)$ denotes the existence of $0 + 2 + 1 = 3$ anti-commuting generators, of which 0 square to $+1$, 2 square to -1 and 1 squares to 0.

[41, Chapter 4]). In this chapter, the vector space in question is $V = \mathbb{F} \times \mathbb{F}$ with $B(v, w) := v \cdot w$. In this situation, $\mathbb{F}$ contains isotropic vectors only if $i := \sqrt{-1} \in \mathbb{F}$. In particular, we note that in a field of characteristic p, there are no isotropic vectors when $p \equiv 3 \pmod 4$.

Given a finite point set A, and we define an oriented segment to be an ordered pair $(a, a') \in A^2$ with length $d(a, a')$. If $d(a, a') = 0$, the segment is called isotropic; it is non-trivial if $a \neq a'$. Any non-trivial isotropic segment lies on an isotropic line with slope $\pm i$.

Isotropic line segments should be excluded from counts: a single isotropic line supporting N points contains $\gg N^2$ zero-length segments. In the context of distinct distances, this is a trivial example which we wish to exclude.

We are now able to define a perpendicular bisector, which is relative to *n*on-isotropic vectors only.

Definition 8.14. *Suppose $u, v \in \mathbb{F}^2$ are such that $(u - v) \cdot (u - v) \neq 0$. Then the perpendicular bisector of u and v is*

$$\mathcal{B}(u, v) := \{x \in \mathbb{F} \times \mathbb{F} : d(u, x) = d(v, x)\}.$$

The set $\mathcal{B}(u, v)$ is a line.

Amongst other facts on isotropic lines, we recall the following fact (stated e.g. in [63]).

Lemma 8.15. *[63, Corollary 8] Perpendicular bisectors are not isotropic lines.*

Proof. Let $u, v \in \mathbb{F}^2$ be such that $(u-v)\cdot(u-v) \neq 0$. Suppose for contradiction that the perpendicular bisector line $\mathcal{B}(u, v)$ is isotropic. Thus every vector in $\mathcal{B}(u, v)$ is isotropic.

Since $d(u, \frac{u-v}{2}) = d(v, \frac{u-v}{2})$, it follows that $\frac{u-v}{2}$ must be an isotropic vector. This is a contradiction since $(u - v) \cdot (u - v) \neq 0$. □

Axial Symmetries

As in the Euclidean case, $\mathrm{SF}_2(\mathbb{F})$ is one of the two cosets in the group of all distance-preserving transformations corresponding to compositions of translations with transformations in $SO_2(\mathbb{F})$. The other coset consists of compositions of an axial symmetry (i.e. reflection) relative to some (non-isotropic) line, and a translation parallel to this line.

As with the matrix interpretation of elements of $\mathrm{SF}_2(\mathbb{F})$ discussed on page 118, there is also an injective group homomorphism from the coset of axial symmetries (where the group operation is composition of maps) into a coset of $\mathsf{SL}_3(\mathbb{F})$ (where the group operation is matrix multiplication), defined by:

$$\left(\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} u & -v \\ v & u \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} s \\ t \end{pmatrix}\right) \mapsto \begin{pmatrix} u & -v & s \\ v & u & t \\ 0 & 0 & 1 \end{pmatrix}, \tag{8.9}$$

where $u^2 + v^2 = -1$.

The matrix (8.9) corresponds to an axial symmetry over the line ℓ where ℓ is defined to be the set

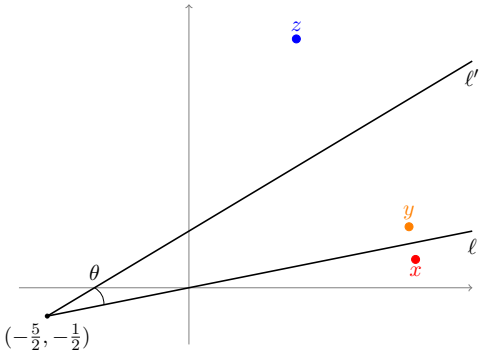
$$\{(a,b) \in \mathbb{F}^2 : : \begin{pmatrix} u & -v \\ v & u \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{pmatrix} a \\ b \end{pmatrix}.$$

Axial symmetries are defined relative to non-isotropic lines only.If a reflected over the line ℓ yields a', then ℓ will be the perpendicular bisector of a and a' so cannot be isotropic.

For $x, y \in \mathbb{F}^2$, we write $x \sim_\ell y$ to mean that x is axially symmetric to y, relative to the (non-isotropic) line ℓ.

The composition of two axial symmetries, relative to distinct lines ℓ and ℓ', as in the Euclidean case, is generally a rotation around the axes intersection point, by twice the angle between the lines. If the lines are parallel, it is a translation in the normal direction (note that ℓ, ℓ' are non-isotropic lines).

Figure 8.5: The composition of two axial symmetries relative to ℓ and ℓ' in the reals: x is reflected over ℓ to obtain y, and y is reflected over ℓ' to obtain z. Alternatively, we could have rotated x by 2θ about the intersection point of ℓ and ℓ' to obtain z.



We would like to work within the group structure of $\mathrm{SF}_2(\mathbb{F})$, rather than its other coset, and so we map the set of all axial symmetries to the group $\mathrm{SF}_2(\mathbb{F})$.

This map is defined as follows. An axial symmetry $\rho \in \mathsf{Isom}(\mathbb{F}^2) \setminus \mathrm{SF}_2(\mathbb{F})$ is mapped to $\mathrm{SF}_2(\mathbb{F})$ by composing it with the fixed axial symmetry relative to the (non-isotropic) x-axis: let r_x be this transformation, so $r_x(a_1, a_2) = (a_1, -a_2)$. Then the axial symmetry associated to ρ is $r_x \circ \rho$.

If we compose the set of all axial symmetries with the symmetry relative to the x-axis, we obtain the set of rotations around all points $(x_0, 0)$ on the x-axis. We call this set $R_x \subseteq \mathrm{SF}_2(\mathbb{F})$.

Explicitly in matrix form:

$$R_x = \left\{ \begin{pmatrix} u & -v & x_0(1-u) \\ v & u & -x_0 v \\ 0 & 0 & 1 \end{pmatrix} : u^2 + v^2 = 1, u, v, x_0 \in \mathbb{F} \right\}.$$

It is then a short calculation to see that the image of R_x under the kinematic mapping lies in the plane $X_2 = 0$. Indeed,

$$\kappa(R_x) = \{[\tilde{u} : \tilde{v} : 0 : x_0\tilde{v}] : (\tilde{u}, \tilde{v}) \in \mathcal{C}, x_0 \in \mathbb{F}\},$$

using the equations (8.6) and (8.7), and the fact that $(\tilde{u}, \tilde{v}) \in \mathcal{C}$. This transformation naturally motivates the role of incidence geometry.

Incidence Geometry

In this chapter, we will use Rudnev's incidence bound [89] between points and planes (Theorem 2.9), Kollár's incidence bound [57] between points and lines in $\mathbb{F}^3$ (Theorem 2.8), and the incidence bound between points and lines in finite fields of de Zeeuw and the author [108] (Theorem 4.5). We restate the latter is its dual form to fit our purposes.

Theorem 8.16 (Points-Lines in $\mathbb{F}^2$)**.** *Let $B \subset \mathbb{F}^*$, consider the set of lines $\mathcal{L}(B)$ indexed by the Cartesian product $B \times B$: a line $\ell_{b,c} \in \mathcal{L}(B)$ is of the form $x \mapsto bx + c$ for $b, c \in B$.*

Let $\mathcal{P} \subset \mathbb{F}^2$ be a finite set of points where $|B| \leq |\mathcal{P}|$. In positive characteristic suppose, in addition, that $|B||\mathcal{P}| \ll p^2$.

Then there exists an absolute constant C such that

$$\mathcal{I}(\mathcal{P}, \mathcal{L}(B)) \leq |\mathcal{P}| + C|B|^{5/4}|\mathcal{P}|^{3/4}. \tag{8.10}$$

We will use Theorem 8.16 to bound the number of k-rich points in a plane. It is an unusual feature of our application that we will already need to have a bound on the number of k-rich points in order to satisfy the constraints of Theorem 8.16. However in our application, because k is large, we can use a trivial estimate (from Lemma 2.1) on the number of k-rich points that is sufficient to satisfy the constraints for Theorem 8.16.

8.8 Proof of Theorems 8.10 and 8.11

Section 8.6, and in particular the high-level overview of Section 8.6 provide a sketch of the proof strategy of Theorems 8.10 and 8.11 which we prove in this section.

Without loss of generality we will assume that at most $c|A|^{2/3}$ points of A lie on a single isotropic line. This is a stronger assumption that in the statement of Theorem 8.10 and Theorem 8.11. We justify this as follows: if some line ℓ contain at least $c|A|^{2/3}$ points of A, but also a point $v \in A$ *not* on ℓ, then there are at least $c|A|^{2/3}/2$ distinct distances of the form $d(v,a)$ for $a \in \ell$. Indeed, a circle centred at v can intersect the line ℓ at most twice.

From pinned distances to isosceles triangles

We first relate the pinned distance problem to that of counting isosceles triangles via the Cauchy-Schwarz inequality. Let $\mathcal{T} = \mathcal{T}(A)$ be the number of non-degenerate isosceles triangles with vertices in A. A non-degenerate isosceles triangle means a triple (a, b, b') with $d(a,b) = d(a,b')$ and the base $b - b'$ non-isotropic.

An application of Cauchy-Schwarz relates $\mathcal{T}$ to the number of pinned distances. This a standard calculation, but we provide a proof here to explain the technical details in an isotropic vector space.

Lemma 8.17. *Let $A \subset \mathbb{F} \times \mathbb{F}$ be a set of cardinality $|A| \geq 2$. Then*

$$\Delta_{pin}(A)\mathcal{T} \gg |A|^3 .$$

Proof. We have

$$\begin{aligned}|A|^2 &= \sum_{r\in\mathbb{F}}\sum_{a,b\in A}\mathbb{1}_{d(a,b)=r}\\ &= \sum_{a\in A}\left(\sum_{r\in\mathbb{F}^*}\sum_{b\in A}\mathbb{1}_{d(a,b)=r}\right)+\sum_{a,b\in A}\mathbb{1}_{d(a,b)=0}\,.\end{aligned}$$

A pair $(a,b)\in A^2$ contributes to the second term if either $a=b$ (of which there are $|A|$ such pairs) or if the vector $a-b$ is isotropic. Since we assume that at most $c|A|^{2/3}$ elements of A lie on an isotropic line, it follows that the second term is bounded by $O(|A|^{5/3})$ and so is indeed an error term.

It remains to bound the first term. We apply the Cauchy-Schwarz inequality in the variable r to obtain:

$$\sum_{a\in A}\left(\sum_{r\in\mathbb{F}^*}\sum_{b\in A}\mathbb{1}_{d(a,b)=r}\right)\leq\sum_{a\in A}\left(\sum_{r\in\mathbb{F}^*}\sum_{b,c\in A}\mathbb{1}_{d(a,b)=d(a,c)=r}\right)^{1/2}|\Delta_a(A)\backslash\{0\}|^{1/2}\,.$$

Recall that $\Delta_a(A)=\{d(a,b)\colon b\in A\}$.

We combine these two equations and perform another application of Cauchy-Schwarz to obtain:

$$\begin{aligned}|A|^2-\binom{|A|}{2} &\leq \left(\sum_{a\in A}\sum_{r\in\mathbb{F}^*}\sum_{b,c\in A}\mathbb{1}_{d(a,b)=d(a,c)=r}\right)^{1/2}\left(\sum_{a\in A}|\Delta_a(A)|\right)^{1/2}\\ &\leq \left(\sum_{r\in\mathbb{F}^*}\sum_{a,b,c\in A}\mathbb{1}_{d(a,b)=d(a,c)=r}\right)^{1/2}|A|^{1/2}\Delta_{\text{pin}}^{1/2}\\ &:= |A|^{1/2}\Delta_{\text{pin}}^{1/2}(\mathcal{T}+\mathcal{T}_{\text{iso}})^{1/2}\,,\end{aligned}$$

where we define

$$\mathcal{T}_{\text{iso}} := |\{(a,b,b')\in A^3 : b-b' \text{ isotropic and } d(a,b)=d(a,b')\neq 0\}|$$

to be the set of isotropic isosceles triangles. As before, $\mathcal{T}=\mathcal{T}(A)$ is the set of non-isotropic isosceles triangles determined by A.

We now have that

$$\frac{|A|(|A|+1)^2}{4}\leq\Delta_{\text{pin}}(\mathcal{T}+\mathcal{T}_{\text{iso}})\,.$$

We claim that $\mathcal{T}_{\text{iso}}\leq 4|A|^2$. This will complete the proof.

To prove this claim, suppose we have a triple $(a, b, c) \in A^3$ contributing to $\mathcal{T}_{\text{iso}}$, and suppose that $b \neq b'$. Suppose we fix $a, b \in A$. Then from the equations $d(b-c, b-c) = 0$ and $d(a, b) = d(a, c)$, we obtain a system of two simultaneous equations, one quadratic and one linear in two unknowns $c = (c_1, c_2)$.

$$c_1(a_1 - b_1) + c_2(a_2 - b_2) = a_1b_1 + a_2b_2 \tag{8.11}$$

$$c_1^2 - 2b_1c_1 + c_2^2 - 2b_2c_1 + b_1^2 + b_2^2 = 0 \tag{8.12}$$

We solve (8.12) in terms of c_1 to find $c_1 = b_1 \pm \frac{1}{2}\sqrt{c_2^2 - 2b_2c_2 + b_2^2 + b_1^2}$. Hence (8.12) becomes a choice of two quadratic equations in c_2. Hence, for fixed a, b, there are at most 4 choices for c. □

From Lemma 8.17, we see that we can translate an upper bound on $\mathcal{T}(A)$ to a lower bound on the number of pinned distances.

Bounding isosceles triangles

We count separately triples of the form (a, a, b) and permutations thereof, and obtain a count for $\mathcal{T}$ in terms of the perpendicular bisectors of pairs of points in A:

$$\mathcal{T} \ll |A|^2 + \sum_{\ell \in \mathcal{B}_1(A)} n_1(\ell) n_2(\ell).$$

In the sum above:

1. $\mathcal{B}_1(A)$ is the set of (non-isotropic) line-bisectors determined by A containing at least one point of A
2. $n_1(\ell) \geq 1$ is the number of points of A lying on ℓ,
3. $n_2(\ell)$ is the number of points $a \in A$ such that $a \notin \ell$ and a has a point in A that is symmetric relative to ℓ.
   i.e. $n_2(\ell) := |\{a \in A \setminus (A \cap \ell) : \exists\ b \in A \text{ such that } a \sim_\ell b\}|$.
   Since perpendicular bisectors are not isotropic, this is well-defined.

We will assume that the sum is over lines $\ell \in \mathcal{B}_1(A)$ for which $n_2(\ell) \gg 1$, with the notation subsuming the count of these lines. Given a (non-isotropic) line bisector ℓ, we may also assume that at least, say, half of the pairs of points axially symmetric with respect to ℓ do not lie on a single isotropic line.

The quantity $\sum_\ell n_1(\ell)^2$ counts, for each line $\ell \in \mathcal{B}_1(A)$, the number of pairs of points lying on ℓ: this is at most $|A|^2$.

Count of isosceles triangles with incidence geometry

We bound $\sum n_2^2(\ell)$ by realising it as an incidence problem. We first introduce some notation.

Let $\mathrm{Ax}_{(c,d)}$ be the set of elements $(x,y) \in \mathbb{F}^2 \times \mathbb{F}^2$ that are axially symmetric to $(c,d) \in \mathbb{F}^2$ (with respect to some non-isotropic line):

$$\mathrm{Ax}_{(c,d)} := \{(x,y) \in \mathbb{F}^2 \times \mathbb{F}^2 \colon \exists \ell \text{ non-isotropic with } (c,d) \sim_\ell (x,y)\}\,.$$

For a non-isotropic set $X \subseteq A \times A$ – that is, for $x = (a,b) \in X$, the vector $a - b$ is non-isotropic – let $\mathcal{A}(X) := \{\mathrm{Ax}_x : x \in X\}$ be the set of sets of elements Ax_x attainable from elements $x \in X$ via axial symmetries. Recall that $S_r \subseteq A^2$ is the set of segments of length r with endpoints in A.

Then:

$$\begin{aligned}\sum_\ell n_2^2(\ell) &= |\{(a,b,c,d,\ell) \in A^4 \times \mathcal{B}_1(A) \colon (a,b) \sim_\ell (c,d)\}| \\ &= \sum_{r \neq 0} |\{((a,b),(c,d) \in S_r^2 : (a,b) \in \mathrm{Ax}_{(c,d)}\}| + |A|^2 \\ &= \sum_{r \neq 0} \mathcal{I}(S_r, \mathcal{A}(S_r)) + |A|^2\,.\end{aligned}$$

The $|A|^2$ summand in the above equation deals with the case $r = 0$. Since the sum over ℓ is over *non-isotropic* lines ℓ, this contribution counts $(a,b) = (c,d)$.

Claim 1. *Let $r \neq 0$, and suppose, if $\mathbb{F}$ has positive characteristic p, that $|A| \leq p^{4/3}$. Suppose that at most M points of A are collinear or co-circular in $\mathbb{F}^2$. Then*

$$\mathcal{I}(S_r, \mathcal{A}(S_r)) \ll M|S_r| + |S_r|^{3/2}\,.$$

Proof. As in the argument above the statement of Proposition 8.12, in this section we suppose that $\mathbb{F}$ is algebraically closed. The set S_r is naturally embedded in $\mathrm{SF}_2(\mathbb{F})$: let s_r be a fixed segment of the form $((0,0),(\rho,0))$, where $\rho^2 = r$. We identify an element $(a,a') \in S_r$ with the inverse of the rigid motion that takes s_r to (a,a'). This rigid motion always exists, for one can translate (a,a') to the origin, and then find the corresponding rotation, for $r \neq 0$.

Once elements of S_r have been identified with points in $\mathrm{SF}_2(\mathbb{F})$, we apply the kinematic mapping κ, to obtain a point $g \in \mathbb{PF}^3$. To summarise, the point g corresponds to a fixed line segment $(a,a') \in A^2$ of length r.

We claim that, for fixed g, the set $\{g \circ \kappa(r) : r \in R_x\}$ lies in a plane (see Section 8.7 for the definition of R_x). Here, $\circ$ is taken to be multiplication of two Clifford algebra elements.

Indeed, as a Clifford algebra, the set $\kappa(R_x)$ is:

$$\kappa(R_x) = \{[1 + u : v : 0 : vx_0] : (u, v) \in \mathcal{C}, x_0 \in \mathbb{F}\}.$$

Then, for a fixed $g = [X_0, X_1, X_2, X_3]$ we have, using the language of Clifford algebras in in Section 8.7, that

$$g \circ R_x = X \circ [\frac{\tilde{u}}{\tilde{v}} : 1 : 0 : x_0] = \frac{\tilde{u}}{\tilde{v}}[X_0 : X_1 : X_2 : X_3] + x_0[0 : 0 : X_1 : X_0] - [X_1 : X_0 : -X_3 : X_2]\,.$$

For further details about the Clifford algebraic set-up of this approach, we refer to reader to the appendix of [73], where a more abstract presentation is provided.

Hence the line segment $(a, a') \in A^2$ is transformed into a set lying on a plane. Moreover, different segments yield different planes.

Furthermore, the endpoints of the set of segments, axially symmetric to two distinct chosen ones (of the same length) lie on a circle or line in $\mathbb{F}^2$ (this calculation is the content of e.g. [63, Lemma 5]). It follows that the quantity M can be used as a bound for the number of collinear planes in the application of the point-plane theorem. Since κ is a projective map, and invoking Theorem 2.9, we have:

$$\mathcal{I}(S_r, \mathcal{A}(S_r)) = \mathcal{I}(|S_r| \text{ points }, |S_r| \text{ planes}) \ll M|S_r| + |S_r|^{3/2}\,,$$

as required. □

Proof of Theorem 8.11

We now conclude the proof of Theorem 8.11. We have (after a further application of Cauchy-Schwarz) that:

$$\begin{aligned}\mathcal{T} &\ll |A|^2 + |A| \left(\sum_{r \neq 0} (M|S_r| + |S_r|^{3/2})\right)^{1/2} \\ &\ll \sqrt{M}|A|^2 + |A|^{3/2} \left(\sum_{r \neq 0} |S_r|^2\right)^{1/4}.\end{aligned}$$

If the first term dominates, then we have $\Omega(M) \gg |A|^{3/4}$ pinned distances. Indeed, if M points of A are collinear, suppose $a_1, \dots, a_M$ are collinear points on a non-isotropic line ℓ. Then $d(a_1, a_i)$ are distinct for at least $(M-1)/2$ values of $i \in \{2, \dots, M\}$. This is because the circle of radius r centred at a_1 intersects ℓ in at most two points. Hence $\Delta_{\text{pin}}(A) \gg M$. On the other hand, if M points of A are cocircular, then we repeat this argument using instead the argument that two distinct circles coincide in at most two points.

Otherwise, the claim of Theorem 8.11 follows from either Proposition 8.12 or the trivial count $\sum_r |S_r|^2 \leq |A|^{7/2}$. This trivial count follows because the maximum realisations of a single non-zero distance is $|A|^{3/2}$ (see e.g. Erdős [29]), and the inequality $\sum_r |S_r| \leq |A|^2$.

Proof of Theorem 8.10

With the tools we have already introduced, we are in fact equipped to prove the stronger Theorem 8.10. We have shown on 128 that

$$\mathcal{T} \ll |A|^2 + |A| \sqrt{M|A|^2 + |A| \left(\sum_{r \neq 0} |S_r|^2 \right)^{1/2}} .$$

As before, if the first term dominates, then $\Delta_{\text{pin}}(A) \gg |A|$ and so we are done.

Suppose instead that the last term dominates, so that $\mathcal{T} \ll |A|^{3/2} (\sum |S_r|^2)^{1/4}$. Instead of bounding $\sum |S_r|^2$ in terms of $|A|$, we can easily relate this quantity to the number of isosceles triangles.

We have

$$\begin{aligned} \sum_{r \neq 0} |S_r|^2 &= \sum_{r \neq 0} \left(\sum_{a \in A} \sum_{b \in A} \mathbb{1}_{d(a,b)=r} \right)^2 \\ &\leq |A| \sum_{r \neq 0} \sum_{a \in A} \sum_{b,b' \in A} \mathbb{1}_{d(a,b)=r} \mathbb{1}_{d(a,b')=r} \\ &= |A| \sum_{a \in A} \sum_{b,b' \in A} \mathbb{1}_{d(a,b)=d(a,b')} \ll |A| \mathcal{T} . \end{aligned}$$

In the above, we used the Cauchy-Schwarz inequality. In this case, we have $\mathcal{T} \ll |A|^{7/4} \mathcal{T}^{1/4}$, from which we deduce that $\mathcal{T} \ll |A|^{7/3}$. Lemma 8.17 concludes the proof of Theorem 8.10.

Finally, suppose that the middle term of (8.8) dominates. Then, on the one hand

$$M|A|^2 \geq |A|(|A|\mathcal{T})^{1/2},$$

using the estimate $\sum |S_r|^2 \ll |A|\mathcal{T}$ (if the middle term is less than this quantity, then by the arguments analysing the case when the third term dominates, we are already done). On the other hand,

$$\mathcal{T} \ll |A|^2 M^{1/2}.$$

Hence

$$\Delta_{\text{pin}}(A) \gg \max(|A|^2 M^{-2}, |A|M^{-1/2}, M).$$

Calculations show that $\Delta_{\text{pin}}(A) \gg |A|^{2/3}$, concluding the proof.

8.9 Proof of Proposition 8.12

In this section we prove the now-redundant Proposition 8.12, which bounds the quantity $\sum_{r\neq 0} |S_r|^2$. This coincides with the notion of "distance energy", or "distance quadruples" in the language of Guth and Katz [43].

By a standard calculation (see e.g. Elekes and Sharir [27, Section **(H3)**]:

$$\sum_{r\neq 0} |S_r|^2 = \sum_{g\in \mathrm{SF}_2(\mathbb{F})} |A\cap gA|^2 = \sum_{k=1}^{|A|} (2k-1)|G_k|,$$

where G_k is the set of *k-rich transformations* of $A\times A$: $g\in \mathrm{SF}_2(\mathbb{F})$ is a k-rich transformation if $|A\cap gA|\geq k$.

We now bound the quantity $|G_k|$. We do this in two different ways: for small k, we map the rich transformations into rich points within a three dimensional line configuration determined by A; for large k we map rich transformations into a two dimensional configuration.

Proposition 8.18. *Suppose that $A\subseteq \mathbb{F}^2$ and in positive characteristic $|A|\leq p$. Then*

$$|G_k(A)| \ll \frac{|A|^3}{k^{3/2}}.$$

Proposition 8.19. *Suppose $\mathbb{F}=\mathbb{F}_p$ and $p\equiv 3 \bmod 4$. If $A\subseteq \mathbb{F}_p^2$ and $|A|\ll p^{10/17}$, then for $k\geq |A|^{4/5}$ we have*

$$|G_k(A)| \ll \frac{|A|^5}{k^4}. \tag{8.13}$$

These bounds, combined with (8.9), immediately imply Proposition 8.12.

Case 1: low multiplicity

Our proof of Proposition 8.18 is similar to the approach of Elekes-Sharir and Guth-Katz (on the technical level we cite the translation to the finite/general field setting by Bennett, Iosevich and Pakianathan [9]). Like the preceding section, this bound is based on the Blaschke-Grünwarld embedding. This time, we restrict the domain so that the image of an element under the mapping lies on a line (rather than on a plane).

Let SF' be the set of transformations in $\mathrm{SF}_2(\mathbb{F})$ that are not pure translations; that is, $SF' = \mathrm{SF}_2(\mathbb{F})\backslash T$.

First, we bound $|G_k \cap T|$: for any $x, y \in A$, there is a *unique* translation sending x to y. Thus, $k|G_k \cap T| \leq |\{(a,b,t) \in A \times A \times (G_k \cap T) \colon a + t = b\}| \leq |A|^2$, and hence $|G_k \cap T| \leq |A|^2/k$.

Now we bound the 'non-trivial' remaining k-rich transformations of A, namely $|G_k \cap SF'|$, by using an incidence bound between points and lines. We reparameterise the set SF' as lines. Using e.g. Bennett et al. [9], we observe that the following map $\phi : \mathbb{F} \setminus \{\pm \mathrm{i}\} \to \mathrm{SO}_2(\mathbb{F})\backslash \mathrm{I}_2$ is a bijection:

$$\phi(r) = \begin{pmatrix} \frac{r^2-1}{r^2+1} & \frac{-2r}{r^2+1} \\ \frac{2r}{r^2+1} & \frac{r^2-1}{r^2+1} \end{pmatrix}.$$

For fixed elements $a = (a_1, a_2), b = (b_1, b_2) \in A$, the set $\{g \in SF' \colon ga \mapsto b\}$ is expressed as the set of matrices of the form:

$$\begin{pmatrix} \frac{r^2-1}{r^2+1} & \frac{-2r}{r^2+1} & \frac{b_1 r^2 + b_1 + 2ra_2 - a_1 r^2 + a_1}{r^2+1} \\ \frac{2r}{r^2+1} & \frac{r^2-1}{r^2+1} & \frac{b_2 r^2 + b_2 - 2ra_1 - a_2 r^2 + a_2}{r^2+1} \\ 0 & 0 & 1 \end{pmatrix}$$

where $r \in \mathbb{F} \setminus \{\pm i\}$ is a parameter. Under the Blaschke-Grünwald kinematic mapping (as in [27], [43]) this set is mapped to the line:

$$\ell_{ab} := \{[0 : 2 : b_2 + a_2 : a_1 + b_1] + r[2 : 0 : b_1 - a_1 : a_2 - b_2] \colon r \in \mathbb{F}\},$$

possibly without two points, corresponding to $r = \pm i$.

An element $g \in SF'$ is in G_k if and only if there are at least k distinct lines ℓ_{ab} indexed by $a, b \in A$ containing the point $\kappa(g)$. This is Lemma 2.6 of [43]. We conclude that $|G_k|$ is precisely the number of k-rich points in the line configuration $\mathcal{L}(A) = \{\ell_{ab} : a, b \in A\}$. (A point x is k-rich with respect to a line configuration $\mathcal{L}$ if $x \in \ell$ for at least k distinct $\ell \in \mathcal{L}$.)

Proof of Proposition 8.18. From the arguments above, it is enough to bound the quantity $|G_k|$, the number of k-rich points in $\mathcal{L} = \mathcal{L}(A)$. For ease of notation, we will write the set of k-rich points as G_k, as well as for the set of k rich transformations: in light of the previous arguments, the cardinality of these two sets are the same.

If we can apply Theorem 2.8 to our situation, then we are done. Indeed, we have $k|G_k| \leq \mathcal{I}(G_k, \mathcal{L}) \ll |\mathcal{L}||G_k|^{1/3} + |G_k|$; we rearrange this to obtain $|G_k| \ll |\mathcal{L}|^{3/2}k^{-3/2} = |A|^3k^{-3/2}$. Our line set $\mathcal{L}(A)$ contains at most $|A|^2$ lines.

It remains to check the assumptions of Kollár's theorem. Note that for distinct elements $a, b, c \in \mathbb{F}^2$, the lines ℓ_{ab} and ℓ_{ac} do not intersect, and so are not coplanar. Thus, at most $\sqrt{|\mathcal{L}|}$ lines are coplanar.

We also need, in positive characteristic, our point set $G_k = G_k(A)$ to have cardinality $|G_k| \ll p^3$. By [57, Corollary 40], if $\mathcal{L}$ is a set of lines in $\mathbb{F}^3$ with $|\mathcal{L}| \leq 2p^2$ in positive characteristic, and at most $\sqrt{|\mathcal{L}|}$ lines of $\mathcal{L}$ are contained in any plane (and at most $2\sqrt{|\mathcal{L}|}$ in any quadric), then the number of points incidence to at least two lines of $\mathcal{L}$ is $\leq 30|\mathcal{L}|^{3/2}$. Since we have $|\mathcal{L}| \leq p^2$, and so $|G_k| \leq 30p^3$, which suffices. □

Case 2: high multiplicity

The idea for the proof of Proposition 8.19 is illustrated by the following example: we can identify the real plane $\mathbb{R}^2$ with the complex numbers $\mathbb{C}$ so that rotations of $\mathbb{R}^2$ correspond to multiplication by complex numbers of norm 1. Thus to bound $|C_t|$ for $\mathbb{F} = \mathbb{R}$, we can view our points as complex numbers and our transformations as certain *affine transformations* of $\mathbb{C}$.

This argument works whenever $\mathbb{R}$ is replaced by a field $\mathbb{F}$ with a non-square element $\alpha \in \mathbb{F}$ and the dot product on $\mathbb{R}^2$ is replaced by a quadratic form $Q((x_1, x_2)) = x_1^2 - \alpha x_2^2$.

The complexification of $\mathbb{F}$ is formalised by the map $\sigma : \mathbb{F}^2 \to \mathbb{F}[\sqrt{\alpha}]$ where α is a non-square:

$$\sigma(x, y) = x + \sqrt{\alpha}y\,.$$

In the case $\mathbb{F} = \mathbb{F}_p$, with $p \equiv 3 \bmod 4$, we take $\alpha = -1$, and write $K = \mathbb{F}[\sqrt{\alpha}]$.

Let $\varphi : \mathrm{SF}_2(\mathbb{F}) \to \mathrm{Aff}(1, K)$ be defined as:

$$\varphi\left(\begin{pmatrix} u & -v & s \\ v & u & t \\ 0 & 0 & 1 \end{pmatrix}\right) = \begin{pmatrix} u + \alpha v & s + \alpha t \\ 0 & 1 \end{pmatrix}.$$

Proof of Proposition 8.19. As in Proposition 8.18, we bound the size of $G_k = \{g \in \mathrm{SF}_2(\mathbb{F}) \colon |A \cap gA| \geq k\}$, for a fixed set $A \subseteq \mathbb{F}^2$.

Using the maps σ and φ defined in the above argument, we set $\tilde{A} = \sigma(A) \subseteq K$ and $\tilde{G}_k = \{g \in \mathrm{Aff}(1, K) \colon |\tilde{A} \cap g\tilde{A}| \geq k\}$. Then, by the equivariance of the map σ, k-rich transformations of A are sent to k-rich transformations of $\tilde{A}$ by the map φ. Hence $|G_k| \leq |\tilde{G}_k|$.

To bound $|\tilde{G}_k|$, we therefore need to bound the number of k-rich lines as to the point set $\tilde{A} \times \tilde{A}$. If $h \in \mathrm{Aff}(1, K)$ is given by the map $x \mapsto mx + b$ with $m \in K^*, b \in K$, then we write $h^* = (m, b)$. Let $H^*_{xy} = \{h^* : hx = y, h \in \mathrm{Aff}(1, K)\}$; H^* is contained in the line $\ell_{xy} = \{(m, b) \in K^2 : y = mx + b\}$.

If $g \in \tilde{G}_k$, then g^* is incident to at least k lines of the form $\ell_{x,y}$ with $x, y \in \tilde{A}$, so $|\tilde{G}_k|$ is at most the number of k-rich points of the set of lines $\ell_{x,y}$ with $x, y \in \tilde{A}$. There are $|\tilde{A}|^2 = |A|^2$ such lines.

We will bound the number of k-rich points $|\tilde{G}_k|$ by Theorem 8.16. Note that our line set $\mathcal{L}(\tilde{A}) := \{\ell_{x,y} \colon x, y \in \tilde{A}\}$ is indexed by a Cartesian product $\tilde{A} \times \tilde{A}$. If $|\tilde{G}_k| \leq |\tilde{A}|$, then we are done, so in the subsequent we suppose that $|\tilde{G}_k| > |\tilde{A}|$. If $|\tilde{A}||\tilde{G}_k| \ll p^2$ then Theorem 8.16 yields

$$|\tilde{G}_k| \ll \frac{|A|^5}{k^4}.$$

It remains therefore to check that $|\tilde{A}||\tilde{G}_k| \ll p^2$. For that we merely use a trivial estimate, for $k \geq |A|^{4/5}$, that

$$|\tilde{G}_k| \ll \frac{|A|^4}{|A|^{8/5}} = |A|^{12/5},$$

which leads to the constraint $|A| \leq p^{10/17}$, as claimed. This concludes the proof of Proposition 8.19 and hence Theorem 8.11. □

8.10 Future work

The inclusion of the subopstimal Theorem 8.11 demonstrates that a better bound on $\sum_r |S_r|^2$ could be a direction towards a stronger pinned distance bound.

The main tool for Theorem 8.10 is Rudnev's points-planes incidence bound. In the situation in which we apply it, we have sets of points with a particular structure; we ask therefore whether a better incidence bound tailored to this situation exists, thus giving a stronger pinned distance bound.

This chapter records the progress made on the pinned distance problem; perhaps there is progress to be made by returning to the non-pinned distinct distance problem with this framework in mind.

Bibliography

[1] Eyal Ackerman. "On the maximum number of edges in topological graphs with no four pairwise crossing edges". In: *Proceedings of the twenty-second annual symposium on Computational geometry*. 2006, pp. 259–263.

[2] Noga Alon. "Eigenvalues, geometric expanders, sorting in rounds, and Ramsey theory". In: *Combinatorica* 6.3 (1986), pp. 207–219.

[3] Antal Balog and Endre Szemerédi. "A statistical theorem of set addition". In: *Combinatorica* 14.3 (1994), pp. 263–268.

[4] Antal Balog and Trevor D Wooley. "A low-energy decomposition theorem." In: *Quarterly Journal of Mathematics* 68.1 (2017), pp. 207–226.

[5] Boaz Barak, Russell Impagliazzo, and Avi Wigderson. "Extracting randomness using few independent sources". In: *SIAM Journal on Computing* 36.4 (2006), pp. 1095–1118.

[6] Abdul Basit and Adam Sheffer. "Incidences with k-non-degenerate sets and their applications". In: *Journal of Computational Geometry* 5.1 (2014), pp. 284–302.

[7] József Beck. "On the lattice property of the plane and some problems of Dirac, Motzkin and Erdős in combinatorial geometry". In: *Combinatorica* 3.3-4 (1983), pp. 281–297.

[8] Michael Bennett, Derrick Hart, Alex Iosevich, Jonathan Pakianathan, and Misha Rudnev. "Group actions and geometric combinatorics in $\mathbb{F}_q^{d}$". In: *Forum Mathematicum* 29.1 (2017), pp. 91–110.

[9] Mike Bennett, Alex Iosevich, and Jonathan Pakianathan. "Three-point configurations determined by subsets of $\mathbb{F}_q^2$ via the Elekes-Sharir Paradigm". In: *Combinatorica* 34.6 (2014), pp. 689–706.

[10] Pierre-Yves Bienvenu, François Hennecart, and Ilya D Shkredov. "A note on the set A (A+ A)". In: *Moscow Journal of Combinatorics and Number Theory* 8.2 (2019), pp. 179–188.

[11] Wilhelm Blaschke. *Euklidische Kinematik und nichteuklidische Geometrie, 1. 2.* Teubner, 1911.

[12] Jean Bourgain. "A modular Szemerédi–Trotter theorem for hyperbolas". In: *Comptes Rendus Mathematique* 350.17-18 (2012), pp. 793–796.

[13] Jean Bourgain. "More on the sum-product phenomenon in prime fields and its applications". In: *International Journal of Number Theory* 1.01 (2005), pp. 1–32.

[14] Jean Bourgain. "Multilinear exponential sums in prime fields under optimal entropy condition on the sources". In: *Geometric and Functional Analysis* 18.5 (2009), pp. 1477–1502.

[15] Jean Bourgain and Mei-Chu Chang. "On the size of k-fold sum and product sets of integers". In: *Journal of the American Mathematical Society* 17.2 (2004), pp. 473–497.

[16] Jean Bourgain and MZ Garaev. "On a variant of sum-product estimates and explicit exponential sum bounds in prime fields". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 146.1 (2009), pp. 1–21.

[17] Jean Bourgain, Nets Katz, and Terence Tao. "A sum-product estimate in finite fields, and applications". In: *Geometric & Functional Analysis GAFA* 14.1 (2004), pp. 27–57.

[18] Oliver Byrne. *The first six books of the Elements of Euclid: in which coloured diagrams and symbols are used instead of letters for the greater ease of learners.* William Pickering, 1847.

[19] Jeremy Chapman, M Burak Erdoğan, Derrick Hart, Alex Iosevich, and Doowon Koh. "Pinned distance sets, k-simplices, Wolff's exponent in finite fields and sum-product estimates". In: *Mathematische Zeitschrift* 271.1-2 (2012), pp. 63–93.

[20] Bernard Chazelle, Herbert Edelsbrunner, and Leonidas J Guibas. "The complexity of cutting complexes". In: *Discrete & Computational Geometry* 4.2 (1989), pp. 139–181.

[21] Fan R. K. Chung. "The number of different distances determined by n points in the plane". In: *Journal of Combinatorial Theory, Series A* 36.3 (1984), pp. 342–354.

[22] Fan RK Chung, Endre Szemerédi, and William T. Trotter. "The number of different distances determined by a set of points in the Euclidean plane". In: *Discrete & Computational Geometry* 7.1 (1992), pp. 1–11.

[23] Kenneth L Clarkson, Herbert Edelsbrunner, Leonidas J Guibas, Micha Sharir, and Emo Welzl. "Combinatorial complexity bounds for arrangements of curves and spheres". In: *Discrete & Computational Geometry* 5.2 (1990), pp. 99–160.

[24] Zeev Dvir. "Incidence theorems and their applications". In: *Foundations and Trends in Theoretical Computer Science* 6.4 (2012), pp. 257–393.

[25] György Elekes. "Sums versus products in number theory, algebra and Erdős geometry". In: *Paul Erdős and his Mathematics II* 11 (2001), pp. 241–290.

[26] György Elekes and Imre Z Ruzsa. "Few sums, many products". In: *Studia Scientiarum Mathematicarum Hungarica* 40.3 (2003), pp. 301–308.

[27] György Elekes and Micha Sharir. "Incidences in three dimensions and distinct distances in the plane". In: *Combinatorics, Probability and Computing* 20.4 (2011), pp. 571–608.

[28] György Elekes and Csaba D Tóth. "Incidences of not-too-degenerate hyperplanes". In: *Computational Geometry (SCG '05)* 6.08 (2005), pp. 16–21.

[29] Paul Erdős. "On sets of distances of n points". In: *The American Mathematical Monthly* 53.5 (1946), pp. 248–250.

[30] Paul Erdős. "On some problems of elementary and combinatorial geometry". In: *Annali di Matematica pura ed applicata* 103.1 (1975), pp. 99–108.

[31] Paul Erdős and Endre Szemerédi. "On sums and products of integers". In: *Studies in pure mathematics*. Springer, 1983, pp. 213–218.

[32] Kenneth J Falconer. "On the Hausdorff dimensions of distance sets". In: *Mathematika* 32.2 (1985), pp. 206–212.

[33] Kevin Ford. "Sums and products from a finite set of real numbers". In: *The Ramanujan Journal* 2.1-2 (1998), pp. 59–66.

[34] Kevin Ford. "The distribution of integers with a divisor in a given interval". In: *Annals of mathematics* (2008), pp. 367–433.

[35] Gregory A Freiman. "The addition of finite sets. I". In: *Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika* 6 (1959), pp. 202–213.

[36] Moubariz Z Garaev. "An Explicit Sum-Product Estimate in $\mathbb{F}_p$". In: *International Mathematics Research Notices* 2007 (2007).

[37] Alexey Glibichuk and Misha Rudnev. "On additive properties of product sets in an arbitrary finite field". In: *Journal d'Analyse Mathématique* 108.1 (2009), p. 159.

[38] William Timothy Gowers. "A new proof of Szemerédi's theorem for arithmetic progressions of length four". In: *Geometric and Functional Analysis* 8.3 (1998), pp. 529–551.

[39] Ben Green and Imre Z Ruzsa. "Freiman's theorem in an arbitrary abelian group". In: *Journal of the London Mathematical Society* 75.1 (2007), pp. 163–175.

[40] Codruţ Grosu. "$\mathbb{F}_p$ is locally like $\mathbb{C}$". In: *Journal of the London Mathematical Society* 89.3 (2014), pp. 724–744.

[41] Larry C Grove. *Classical groups and geometric algebra*. Vol. 39. American Mathematical Soc., 2002.

[42] Josef Grünwald. "Ein Abbildungsprinzip, welches die ebene Geometrie und Kinematik mit der räumlichen Geometrie verknüpft". In: *Sitzber. Ak. Wiss. Wien* 120 (1911), pp. 677–741.

[43] Larry Guth and Nets Hawk Katz. "On the Erdős distinct distances problem in the plane". In: *Annals of Mathematics* (2015), pp. 155–190.

[44] Brandon Hanson, Ben Lund, and Oliver Roche-Newton. "On distinct perpendicular bisectors and pinned distances in finite fields". In: *Finite Fields and Their Applications* 37 (2016), pp. 240–264.

[45] Derrick Hart and Alex Iosevich. "Sums and products in finite fields: an integral geometric viewpoint". In: *Radon transforms, geometry, and wavelets* 464 (2008), pp. 129–135.

[46] Harald Andrés Helfgott and Misha Rudnev. "An explicit incidence theorem in $\mathbb{F}_p$". In: *Mathematika* 57.1 (2011), pp. 135–145.

[47] Alex Iosevich, Doowon Koh, and Thang Pham. "A new perspective on the distance problem over prime fields". In: *arXiv preprint arXiv:1905.04179* (2019).

[48] Alex Iosevich, Doowon Koh, Thang Pham, Chun-Yen Shen, and Le Anh Vinh. "A new bound on Erdős distinct distances problem in the plane over prime fields". In: *arXiv preprint arXiv:1805.08900* (2018).

[49] Alex Iosevich and Misha Rudnev. "Erdős distance problem in vector spaces over finite fields". In: *Transactions of the American Mathematical Society* 359.12 (2007), pp. 6127–6142.

[50] Timothy GF Jones. "An improved incidence bound for fields of prime order". In: *European Journal of Combinatorics* 52 (2016), pp. 136–145.

[51] Timothy GF Jones. "Further improvements to incidence and Beck-type bounds over prime finite fields". In: *arXiv preprint arXiv:1206.4517* (2012).

[52] Timothy GF Jones. "New quantitative estimates on the incidence geometry and growth of finite sets". In: *arXiv preprint arXiv:1301.4853* (2013).

[53] Nets Hawk Katz. "An improvement of a lemma of Tardos". In: *unpublished* (2003). [Online; accessed 2019-11-21]. URL: `%5Curl%7Bhttps://kam.mff.cuni.cz/~matousek/cla/katz-betterdistdist%7D`.

[54] Nets Hawk Katz and Chun-Yen Shen. "A slight improvement to Garaev's sum product estimate". In: *Proceedings of the American Mathematical Society* 136.7 (2008), pp. 2499–2504.

[55] Nets Hawk Katz and Gábor Tardos. "A new entropy inequality for the Erdős distance problem". In: *Contemporary Mathematics* 342 (2004), pp. 119–126.

[56] Doowon Koh, Thang Pham, and Le Anh Vinh. "Extension theorems and Distance problems over finite fields". In: *arXiv preprint arXiv:1809.08699v8* (2018).

[57] János Kollár. "Szemerédi–Trotter-type theorems in dimension 3". In: *Advances in Mathematics* 271 (2015), pp. 30–61.

[58] Sergei V Konyagin and Ilya D Shkredov. "New results on sums and products in $\mathbb{R}$". In: *Proceedings of the Steklov Institute of Mathematics* 294.1 (2016), pp. 78–88.

[59] Sergei Vladimirovich Konyagin. "A sum-product estimate in fields of prime order". In: *arXiv preprint math/0304217* (2003).

[60] Sergei Vladimirovich Konyagin and Ilya D Shkredov. "On sum sets of sets having small product set". In: *Proceedings of the Steklov Institute of Mathematics* 290.1 (2015), pp. 288–299.

[61] Mark Lewko. "Counting rectangles and an improved restriction estimate for the paraboloid in $\mathbb{F}_p^3$". In: *Proceedings of the American Mathematical Society* (2020).

[62] Liangpan Li. "Slightly improved sum-product estimates in fields of prime order". In: *Acta Arithmetica* 147.2 (2011), pp. 153–160.

[63] Ben Lund and Giorgis Petridis. "Bisectors and pinned distances". In: *Discrete & Computational Geometry* (2019), pp. 1–18.

[64] Ben Lund, Adam Sheffer, and Frank De Zeeuw. "Bisector energy and few distinct distances". In: *Discrete & Computational Geometry* 56.2 (2016), pp. 337–356.

[65] Jiří Matoušek. *Lectures on discrete geometry*. Vol. 108. Springer, 2002.

[66] Ali Mohammadi. "Szemerédi-Trotter type results in arbitrary finite fields". In: *arXiv preprint arXiv:1808.05543* (2018).

[67] Leo Moser. "On the different distances determined by n points". In: *The American Mathematical Monthly* 59.2 (1952), pp. 85–91.

[68] Brendan Murphy and Giorgis Petridis. "A point-line incidence identity in finite fields, and applications". In: *Moscow Journal of Combinatorics and Number Theory* 6.1 (2016), pp. 63–94.

[69] Brendan Murphy and Giorgis Petridis. "An example related to the Erdős -Falconer question over arbitrary finite fields". In: *arXiv preprint arXiv:1905.05634* (2019).

[70] Brendan Murphy, Giorgis Petridis, Oliver Roche-Newton, Misha Rudnev, and Ilya D Shkredov. "New results on sum-product type growth over fields". In: *Mathematika* 65.3 (2019), pp. 588–642.

[71] Brendan Murphy, Oliver Roche-Newton, and Ilya D Shkredov. "Variations on the sum-product problem II". In: *SIAM Journal on Discrete Mathematics* 31.3 (2017), pp. 1878–1894.

[72] Brendan Murphy, Misha Rudnev, Ilya D Shkredov, and Yurii N Shteinikov. "On the few products, many sums problem". In: *arXiv preprint arXiv:1712.00410* (2017).

[73] Brendan Murphy, Misha Rudnev, and Sophie Stevens. "Bisector energy and pinned distances in positive characteristic". In: *arXiv preprint arXiv:1908.04618* (2019).

[74] Melvyn Nathanson. "On sums and products of integers". In: *Proceedings of the American Mathematical Society* 125.1 (1997), pp. 9–16.

[75] Melvyn B. Nathanson. *Additive Number Theory: Inverse Problems and the Geometry of Sumsets*. Springer Science & Business Media, 1996.

[76] Melvyn B Nathanson and Gérald Tenenbaum. "Inverse theorems and the number of sums and products". In: *Astérisque* 258 (1999), pp. 195–204.

[77] Konstantin I. Olmezov, Aliaksei S. Semchankau, and Ilya D. Shkredov. "On popular sums and differences of sets with small products". In: *arXiv preprint arXiv:1911.12005* (2019).

[78] János Pach and Pankaj K Agarwal. *Combinatorial geometry*. John Wiley & Sons, 1995.

[79] Giorgis Petridis. "Collinear triples and quadruples for Cartesian products in $\mathbb{F}_p^2$ ". In: *arXiv preprint arXiv:1610.05620* (2016).

[80] Giorgis Petridis. "New proofs of Plünnecke-type estimates for product sets in groups". In: *Combinatorica* 32.6 (2012), pp. 721–733.

[81] Giorgis Petridis. "Pinned algebraic distances determined by Cartesian products in $\mathbb{F}_p^2$ ". In: *Proceedings of the American Mathematical Society* 145.11 (2017), pp. 4639–4645.

[82] Giorgis Petridis. "Products of Differences in Prime Order Finite Fields". In: (2016). arXiv: `1602.02142 [math.CO]`.

[83] Thang Pham, Le Anh Vinh, Frank de Zeeuw, et al. "Three-variable expanding polynomials and higher-dimensional distinct distances". In: *Combinatorica* 39.2 (2019), pp. 411–426.

[84] Helmut Plünnecke. "Eine zahlentheoretische Anwendung der Graphentheorie." In: *Journal für die reine und angewandte Mathematik* 243 (1970), pp. 171–183.

[85] Jürgen Richter-Gebert. *Perspectives on projective geometry: A guided tour through real and complex geometry*. Springer Science & Business Media, 2011.

[86] Oliver Roche-Newton, Misha Rudnev, and Ilya D Shkredov. "New sum-product type estimates over finite fields". In: *Advances in Mathematics* 293 (2016), pp. 589–605.

[87] Oliver Roche-Newton and Audie Warren. "New Expander Bounds from Affine Group Energy". In: *arXiv preprint arXiv:1905.03701* (2019).

[88] Misha Rudnev. "An improved sum–product inequality in fields of prime order". In: *International Mathematics Research Notices* 2012.16 (2012), pp. 3693–3705.

[89] Misha Rudnev. "On the number of incidences between points and planes in three dimensions". In: *Combinatorica* 38.1 (2018), pp. 219–254.

[90] Misha Rudnev. "Point-plane incidences and some applications in positive characteristic". In: vol. 23. Walter de Gruyter GmbH & Co KG, 2019, pp. 211–240.

[91] Misha Rudnev, George Shakan, and Ilya D Shkredov. "Stronger sum-product inequalities for small sets". In: *Proceedings of the AMS* (2019).

[92] Misha Rudnev, Ilya D Shkredov, and Sophie Stevens. "On the energy variant of the sum-product conjecture". In: *arXiv preprint arXiv:1607.05053* (2016).

[93] Imre Z Ruzsa. "An application of graph theory to additive number theory". In: *Scientia, Ser. A* 3.9 (1989), pp. 97–109.

[94] Belkacem Said-Houari. *Linear Algebra*. Springer, 2017.

[95] Tomasz Schoen. "New bounds in Balog-Szemerédi-Gowers theorem". In: *Combinatorica* 35.6 (2015), pp. 695–701.

[96] Tomasz Schoen and Ilya D Shkredov. "Higher moments of convolutions". In: *Journal of Number Theory* 133.5 (2013), pp. 1693–1737.

[97] George Shakan. "On higher energy decompositions and the sum–product phenomenon". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 167.3 (2019), pp. 599–617.

[98] Adam Sheffer. "Incidence Theory". In: *Unfinished manuscript* (). URL: `http://faculty.baruch.cuny.edu/ASheffer/000book.pd`.

[99] Adam Sheffer, Endre Szabó, and Joshua Zahl. "Point-curve incidences in the complex plane". In: *Combinatorica* 38.2 (2018), pp. 487–499.

[100] Ilya D Shkredov. "Modular hyperbolas and bilinear forms of Kloosterman sums". In: *arXiv preprint arXiv:1905.00291* (2019).

[101] Ilya D Shkredov. "On sums of Szemerédi-Trotter sets". In: *Proceedings of the Steklov Institute of Mathematics* 289.1 (2015), pp. 300–309.

[102] Ilya D Shkredov. "Some applications of W. Rudin's inequality to problems of combinatorial number theory". In: *Uniform Distribution Theory* 6.2 (2011), pp. 95–116.

[103] Ilya D Shkredov. "Some new results on higher energies". In: *Transactions of the Moscow Mathematical Society* 74 (2013), pp. 31–63.

[104] József Solymosi. "Bounding multiplicative energy by the sumset". In: *Advances in mathematics* 222.2 (2009), pp. 402–408.

[105] József Solymosi. "On the number of sums and products". In: *Bulletin of the London Mathematical Society* 37.4 (2005), pp. 491–494.

[106] Jozsef Solymosi and Gabor Tardos. "On the number of k-rich transformations". In: *23rd Annual Symposium on Computational Geometry, SCG'07* (2007), pp. 227–231.

[107] József Solymosi and Csaba D Tóth. "Distinct distances in homogeneous sets in Euclidean space". In: *Discrete & Computational Geometry* 35.4 (2006), pp. 537–549.

[108] Sophie Stevens and Frank De Zeeuw. "An improved point-line incidence bound over arbitrary fields". In: *Bulletin of the London Mathematical Society* 49.5 (2017), pp. 842–858.

[109] László A Székely. "Crossing numbers and hard Erdős problems in discrete geometry". In: *Combinatorics, Probability and Computing* 6.3 (1997), pp. 353–358.

[110] Endre Szemerédi and William T. Trotter. "Extremal problems in discrete geometry". In: *Combinatorica* 3.3-4 (1983), pp. 381–392.

[111] Terence Tao and Van H Vu. *Additive combinatorics*. Vol. 105. Cambridge studies in advanced mathematics. Cambridge University Press, 2006.

[112] Terry Tao. *The Szemerédi-Trotter theorem and the cell decomposition*. [Online; accessed 2019-10-22]. 2009. URL: `https://terrytao.wordpress.com/2011/02/18/the-szemeredi-trotter-theorem-via-the-polynomial-ham-sandwich-theorem/`.

[113] Gábor Tardos. "On distinct sums and distinct distances". In: *Advances in Mathematics* 180.1 (2003), pp. 275–289.

[114] Csaba D Tóth. "The Szemerédi–Trotter theorem in the complex plane". In: *Combinatorica* 35.1 (2015), pp. 95–126.

[115] Le Anh Vinh. "On four-variable expanders in finite fields". In: *SIAM Journal on Discrete Mathematics* 27.4 (2013), pp. 2038–2048.

[116] Le Anh Vinh. "The Szemerédi–Trotter type theorem and the sum-product estimate in finite fields". In: *European Journal of Combinatorics* 32.8 (2011), pp. 1177–1181.

[117] Van H Vu, Melanie Matchett Wood, and Philip Matchett Wood. "Mapping incidences". In: *Journal of the London Mathematical Society* 84.2 (2011), pp. 433–445.

[118] Esen Aksoy Yazici, Brendan Murphy, Misha Rudnev, and Ilya D Shkredov. "Growth estimates in positive characteristic via collisions". In: *International Mathematics Research Notices* 2017.23 (2016), pp. 7148–7189.

[119] Joshua Zahl. "A Szemerédi–Trotter Type Theorem in $\mathbb{R}^4$". In: *Discrete & Computational Geometry* 54.3 (2015), pp. 513–572.

[120] Frank de Zeeuw. "A short proof of Rudnev's point-plane incidence bound". In: *arXiv preprint arXiv:1612.02719* (2016).